

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 13:52:29 ; Search time 3483 Seconds
(without alignments)
7167.840 Million cell updates/sec

Title: US-09-975-856-1
Perfect score: 576
Sequence: 1 ATGAACGGAGACGACGCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	576	100.0	576	6	AR167456	AR167456 Sequence
2	576	100.0	576	6	AR170503	AR170503 Sequence
3	576	100.0	576	6	AR182977	AR182977 Sequence
4	576	100.0	576	6	AR305671	AR305671 Sequence
5	576	100.0	576	6	AX740218	AX740218 Sequence
6	576	100.0	576	6	BD096872	BD096872 Isolated
7	576	100.0	576	6	BD218420	BD218420 SSX gene,
8	576	100.0	576	9	HSU90841	U90841 Homo sapien
9	574.4	99.7	1250	9	BC005325	BC005325 Homo sapi
10	496.6	86.2	766	6	A48452	A48452 Sequence 17
11	496.6	86.2	766	6	AX821947	AX821947 Sequence
12	496.6	86.2	766	9	HSSX2	X86175 H. sapiens m
13	496.6	86.2	931	6	AR025465	AR025465 Sequence
14	496.6	86.2	931	6	AR060380	AR060380 Sequence
15	496.6	86.2	931	6	AR117874	AR117874 Sequence
16	496.6	86.2	931	6	AR167453	AR167453 Sequence
17	496.6	86.2	931	6	I85576	I85576 Sequence 2
18	496.6	86.2	931	6	AR287596	AR287596 Sequence
19	496.6	86.2	931	6	AR305668	AR305668 Sequence
20	496.6	86.2	931	6	AX114023	AX114023 Sequence
21	496.6	86.2	931	6	AX719099	AX719099 Sequence
22	496.6	86.2	931	6	BD218417	BD218417 SSX gene,
23	496.6	86.2	1309	6	AX331558	AX331558 Sequence
24	496.6	86.2	1309	6	AX331950	AX331950 Sequence
25	496.6	86.2	1309	9	HSBD21LEK	Z49105 H. sapiens H
26	496	86.1	1249	9	HSU90840	U90840 Homo sapien
27	495	85.9	1276	9	BC016957	BC016957 Homo sapi
28	493.4	85.7	1282	9	BC007343	BC007343 Homo sapi
29	492.8	85.6	675	9	S82471	S82471 Homo sapien
30	482.2	83.7	576	6	AR167457	AR167457 Sequence
31	482.2	83.7	576	6	AR170504	AR170504 Sequence
32	482.2	83.7	576	6	AR182978	AR182978 Sequence
33	482.2	83.7	576	6	AR305672	AR305672 Sequence
34	482.2	83.7	576	6	AX740219	AX740219 Sequence
35	482.2	83.7	576	6	BD096873	BD096873 Isolated
36	482.2	83.7	576	6	BD218421	BD218421 SSX gene,
37	481.6	83.6	1261	9	BC001003	BC001003 Homo sapi
38	481.2	83.5	766	6	A48450	A48450 Sequence 15
39	480.6	83.4	576	9	HSU90842	U90842 Homo sapien
40	480	83.3	766	6	AR167452	AR167452 Sequence
41	480	83.3	766	6	AR305667	AR305667 Sequence
42	480	83.3	766	6	AX821946	AX821946 Sequence
43	480	83.3	766	6	BD218416	BD218416 SSX gene,
44	480	83.3	766	9	HSSX1	X86174 H. sapiens m
45	433.2	75.2	1363	9	BC016640	BC016640 Homo sapi

ALIGNMENTS

RESULT 1
AR167456
LOCUS AR167456
DEFINITION Sequence 5 from patent US 6287756.
ACCESSION AR167456
VERSION AR167456.1 GI:17903237
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 576)
Tureci, O., Chen, Y.-T., Sahin, U., Gure, A.O., Old, L.J. and
Pfreundschuh, M.
AUTHORS
TITLE
Methods for determining presence of cancer in a sample by
determining expression of an SSX gene

AR167456 576 bp DNA linear PAT 17-DEC-2001

JOURNAL Patent: US 6287756-A 5 11-SEP-2001;
FEATURES Location/Qualifiers
source 1..576
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAACGGAGACGCGCTTTGCAAGGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
Db 1 ATGAACGGAGACGCGCTTTGCAAGGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
Qy 61 TTACGAAAGCGCTTCGATGATATGCGCAAAATCTCTTAAGAAAGAGTGCGGAAAGATG 120
Db 61 TTACGAAAGCGCTTCGATGATATGCGCAAAATCTCTTAAGAAAGAGTGCGGAAAGATG 120
Qy 121 AAATCCTCGGAGAAATCGTCTATGTATATGAGCTAAACTATGAGGTCATGACTAAA 180
Db 121 AAATCCTCGGAGAAATCGTCTATGTATATGAGCTAAACTATGAGGTCATGACTAAA 180
Qy 181 CTAGGTTTCAGGTCACCGCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAGGTCACCGCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Qy 241 GGGAAATGATTTGGTAAACGATCGAAACCAACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Db 241 GGGAAATGATTTGGTAAACGATCGAAACCAACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Qy 301 TTGCGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCGCAAGAAATGATGGGAAACAGCTGTGC 420
Db 301 TTGCGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCGCAAGAAATGATGGGAAACAGCTGTGC 420
Qy 361 AATGTTTGAAGGAGTCCAGAGGATCTCGGCCACAAATGATGGGAAACAGCTGTGC 480
Db 361 AATGTTTGAAGGAGTCCAGAGGATCTCGGCCACAAATGATGGGAAACAGCTGTGC 480
Qy 421 CCCCCGGGAAATCCAAAGTACCTTGAGAAAGATTAAACAGACATCTGGACCCCAAAAGGGGG 540
Db 421 CCCCCGGGAAATCCAAAGTACCTTGAGAAAGATTAAACAGACATCTGGACCCCAAAAGGGGG 540
Qy 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576
Db 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576

RESULT 2
LOCUS AR170503 576 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 1 from patent US 6291658.
ACCESSION AR170503
VERSION AR170503.1 GI:17908462
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 576)
AUTHORS Gure,A.O., Tureci,O., Sahin,U., Tsang,S., Scanlan,M.J., Knuth,A., Pfeundscher,M., Old,L.J. and Chen,Y.-T.
TITLE Isolated nucleic acid molecules encoding SSX family members and thereof
JOURNAL Patent: US 6291658-A 1 18-SEP-2001;
FEATURES Location/Qualifiers
source 1..576
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAACGGAGACGCGCTTTGCAAGGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
Db 1 ATGAACGGAGACGCGCTTTGCAAGGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
Qy 61 TTACGAAAGCGCTTCGATGATATGCGCAAAATCTCTTAAGAAAGAGTGCGGAAAGATG 120
Db 61 TTACGAAAGCGCTTCGATGATATGCGCAAAATCTCTTAAGAAAGAGTGCGGAAAGATG 120
Qy 121 AAATCCTCGGAGAAATCGTCTATGTATATGAGCTAAACTATGAGGTCATGACTAAA 180
Db 121 AAATCCTCGGAGAAATCGTCTATGTATATGAGCTAAACTATGAGGTCATGACTAAA 180
Qy 181 CTAGGTTTCAGGTCACCGCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAGGTCACCGCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Qy 241 GGGAAATGATTTGGTAAACGATCGAAACCAACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Db 241 GGGAAATGATTTGGTAAACGATCGAAACCAACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Qy 301 TTGCGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCGCAAGAAATGATGGGAAACAGCTGTGC 420
Db 301 TTGCGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCGCAAGAAATGATGGGAAACAGCTGTGC 420
Qy 361 AATGTTTGAAGGAGTCCAGAGGATCTCGGCCACAAATGATGGGAAACAGCTGTGC 480
Db 361 AATGTTTGAAGGAGTCCAGAGGATCTCGGCCACAAATGATGGGAAACAGCTGTGC 480
Qy 421 CCCCCGGGAAATCCAAAGTACCTTGAGAAAGATTAAACAGACATCTGGACCCCAAAAGGGGG 540
Db 421 CCCCCGGGAAATCCAAAGTACCTTGAGAAAGATTAAACAGACATCTGGACCCCAAAAGGGGG 540
Qy 481 AAACATGCTGGACCCACAGATCGTGGGAAAGAGAGAGAGATC 540
Db 481 AAACATGCTGGACCCACAGATCGTGGGAAAGAGAGAGATC 540
Qy 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576
Db 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576

RESULT 3

LOCUS AR182977 576 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6339140.
ACCESSION AR182977
VERSION AR182977.1 GI:20226184
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 576)
AUTHORS Gure,A.O., Tureci,O., Sahin,U., Tsang,S., Scanlan,M.J., Knuth,A., Pfeundscher,M., Old,L.J. and Chen,Y.-T.
TITLE SSX family proteins
JOURNAL Patent: US 6339140-A 1 15-JAN-2002;
FEATURES Location/Qualifiers
source 1..576
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAACGGAGACGCGCTTTGCAAGGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60

Db 1 ATGACGGAGACGACGCCCTTTGCAAGGAGAGACCCAGGAGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAAGCCCTTCGATGATATTTGCCAAATATCTCTCTAAGAAAGAGTGGGAAAGATG 120
Db 61 TTACGAAAGCCCTTCGATGATATTTGCCAAATATCTCTCTAAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCCTCGAGAAAATCGTCTATGTGTATATGAAGCTAAACTATGAGGTCATGACTTAA 180
Db 121 AAATCCTCGAGAAAATCGTCTATGTGTATATGAAGCTAAACTATGAGGTCATGACTTAA 180
QY 181 CTAGGTTTCAAGTCAACCTCCACACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGTCAACCTCCACACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGAATGATTTTGGTAACGATCGAAACACACAGGAATCAGGTTTGAACCTCTCAGATGACT 300
Db 241 GGAATGATTTTGGTAACGATCGAAACACACAGGAATCAGGTTTGAACCTCTCAGATGACT 300
QY 301 TTCCGAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA 360
Db 301 TTCCGAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA 360
QY 361 AATGGTTTGAAGAGTGCAGAGGATCTGGCCACAAATGATGGAAACAGCTGTGC 420
Db 361 AATGGTTTGAAGAGTGCAGAGGATCTGGCCACAAATGATGGAAACAGCTGTGC 420
QY 421 CCCCCGGGAAATCCAGTACTTGGAGAAGATTAAACAAGACATCTGGACCCCAAAAGGGGG 480
Db 421 CCCCCGGGAAATCCAGTACTTGGAGAAGATTAAACAAGACATCTGGACCCCAAAAGGGGG 480
QY 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
QY 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576
Db 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576

RESULT 4
AR305671
LOCUS AR305671 576 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 5 from patent US 6548064.
ACCESSION AR305671
VERSION AR305671.1 GI:31695155
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 576)
AUTHORS Tureci,O., Sahin,U., Pfreundschuh,M., Rammensee,H.G. and Stevanovic,S.
TITLE Isolated peptides consisting of amino acid sequences found in SSX or NY-ESO-1 molecules, which bind to HLA molecules
JOURNAL Patent: US 6548064-A 5 15-APR-2003;
FEATURES
source Location/Qualifiers
1..576
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACGGAGACGACGCCCTTTGCAAGGAGAGACCCAGGAGATGATGCTCAAAATATCAGAGAAG 60
Db 1 ATGACGGAGACGACGCCCTTTGCAAGGAGAGACCCAGGAGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAAGCCCTTCGATGATATTTGCCAAATATCTCTCTAAGAAAGAGTGGGAAAGATG 120
Db 61 TTACGAAAGCCCTTCGATGATATTTGCCAAATATCTCTCTAAGAAAGAGTGGGAAAGATG 120

QY 121 AAATCCTCGAGAAAATCGTCTATGTGTATATGAAGCTAAACTATGAGGTCATGACTTAA 180
Db 121 AAATCCTCGAGAAAATCGTCTATGTGTATATGAAGCTAAACTATGAGGTCATGACTTAA 180
QY 181 CTAGGTTTCAAGTCAACCTCCACACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGTCAACCTCCACACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGAATGATTTTGGTAACGATCGAAACACACAGGAATCAGGTTTGAACCTCTCAGATGACT 300
Db 241 GGAATGATTTTGGTAACGATCGAAACACACAGGAATCAGGTTTGAACCTCTCAGATGACT 300
QY 301 TTCCGAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA 360
Db 301 TTCCGAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA 360
QY 361 AATGGTTTGAAGAGTGCAGAGGATCTGGCCACAAATGATGGAAACAGCTGTGC 420
Db 361 AATGGTTTGAAGAGTGCAGAGGATCTGGCCACAAATGATGGAAACAGCTGTGC 420
QY 421 CCCCCGGGAAATCCAGTACTTGGAGAAGATTAAACAAGACATCTGGACCCCAAAAGGGGG 480
Db 421 CCCCCGGGAAATCCAGTACTTGGAGAAGATTAAACAAGACATCTGGACCCCAAAAGGGGG 480
QY 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
QY 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576
Db 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576

RESULT 5
AX740218
LOCUS AX740218 576 bp mRNA linear PAT 08-MAY-2003
DEFINITION Sequence 1 from Patent EP1300463.
ACCESSION AX740218
VERSION AX740218.1 GI:30519358
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Isolated nucleic acid molecules encoding SSX family members and uses thereof
JOURNAL Patent: EP 1300463-A 1 09-APR-2003;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US) ; MEMORIAL SLOAN-KETTERING CANCER CENTER (US) ; CORNELL RESEARCH FOUNDATION (US)

FEATURES
source Location/Qualifiers
1..576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene 1..576
/gene="SSX4"
CDS 1..567
/gene="SSX4"
/codon_start=1
/protein_id="CAD90570.1"
/db_xref="GI:30519359"
/translation="MNGDDAFARRPRDDAQISEKLKAFDDIAKIFSKWEKMKSSSE KIVYVYMLNIEVMTKLGKVLTPPMRSKRAADFGNDNRHNRQVRFQMTFG SLQRIPEKTKPEEENGLKEVPEASGPQNDKQLCPPGNPSTLEKINKTSGPKRG KHAWHLRERKQLVYVEISDPEEDE"

ORIGIN
Query Match 100.0%; Score 576; DB 6; Length 576;

Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGAGGAGCGCCCTTTGCAAGGAGAGCCAGGATGATGCTCAATATCAGAGAAG 60
Db 1 ATGACGAGGAGCGCCCTTTGCAAGGAGAGCCAGGATGATGCTCAATATCAGAGAAG 60

QY 61 TTACGAAAGGCGCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
Db 61 TTACGAAAGGCGCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120

QY 121 AAATCCTCGGAGAAATGCTATGTATATGAAGCTAACTATGAGTGTATGACTAAA 180
Db 121 AAATCCTCGGAGAAATGCTATGTATATGAAGCTAACTATGAGTGTATGACTAAA 180

QY 181 CTAGGTTTCAAGTCAAGCTCCACCTTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGTCAAGCTCCACCTTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 240

QY 241 GCGAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGCTTCAGATGACT 300
Db 241 GCGAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGCTTCAGATGACT 300

QY 301 TTCGACAGCTCCAGAGATCTTCCGAGATCATGCCCCAGAGACCCAGCAGAGAAGAA 360
Db 301 TTCGACAGCTCCAGAGATCTTCCGAGATCATGCCCCAGAGACCCAGCAGAGAAGAA 360

QY 361 AATGTTTGAAGGAAGTGCAGAGCATCTGCCCCACAAAATGATGGGAAACAGCTGTGC 420
Db 361 AATGTTTGAAGGAAGTGCAGAGCATCTGCCCCACAAAATGATGGGAAACAGCTGTGC 420

QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCAAAGGGGG 480
Db 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCAAAGGGGG 480

QY 481 AAACATGCTGGACCCAGAGCTGCTGAGAGAGAGCTGCTGTTTATGAAGATC 540
Db 481 AAACATGCTGGACCCAGAGCTGCTGAGAGAGAGCTGCTGTTTATGAAGATC 540

QY 541 AGCGACCTTGAGGAAGATGACGAGTAATCCCTCTCG 576
Db 541 AGCGACCTTGAGGAAGATGACGAGTAATCCCTCTCG 576

BD096872 576 bp DNA linear PAT 27-AUG-2002
LOCUS Isolated nucleic acid molecules encoding SSX family members and uses thereof.
DEFINITION BD096872
ACCESSION BD096872.1 GI:22642460
VERSION JP 2001527408-A/1.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 576)
AUTHORS Gure,A.O., Tureci,O., Sahin,U., Tsang,S., Scanlan,M.J., Knuth,A., Pfreundschuh,M., Old,L.J. and Chen,Y.T.
TITLE Isolated nucleic acid molecules encoding SSX family members and uses thereof
JOURNAL Patent: JP 2001527408-A 1 25-DEC-2001;
LUDWIG INSTITUTE FOR CANCER RESEARCH, MEMORIAL SLOAN KETTERING
CANCER CENTER,CORNELL RESERCH FOUNDATION
COMMENT OS Unidentified
PN JP 2001527408-A/1
PD 25-DEC-2001
PP 25-FEB-1998 JP 1998548050
PR 05-MAY-1997 US 08/851138
PI ALI O GURE,OZLEM TURECI,UGUR SAHIN,SOLAM TSANG,MATTHEW J PI SCANLAN,
PI ALEXANDER KNUTH,MICHAEL PFEUNDSCUH,LLOYD J OLD,YAO TSENG PI CHEN
PC C12N5/10,C12N15/12,C12P21/02,C12Q1/68

CC Strandedness: Single;
CC Topology: Linear;
CC Isolated nucleic acid molecules encoding SSX family members and uses
CC thereof
FH Key Location/Qualifiers
FT source 1..576
FEATURES
source Location/Qualifiers
1..576
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGAGGAGCGCCCTTTGCAAGGAGAGCCAGGATGATGCTCAATATCAGAGAAG 60
Db 1 ATGACGAGGAGCGCCCTTTGCAAGGAGAGCCAGGATGATGCTCAATATCAGAGAAG 60

QY 61 TTACGAAAGGCGCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
Db 61 TTACGAAAGGCGCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120

QY 121 AAATCCTCGGAGAAATGCTATGTATATGAAGCTAACTATGAGTGTATGACTAAA 180
Db 121 AAATCCTCGGAGAAATGCTATGTATATGAAGCTAACTATGAGTGTATGACTAAA 180

QY 181 CTAGGTTTCAAGTCAAGCTCCACCTTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGTCAAGCTCCACCTTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 240

QY 241 GCGAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGCTTCAGATGACT 300
Db 241 GCGAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGCTTCAGATGACT 300

QY 301 TTCGACAGCTCCAGAGATCTTCCGAGATCATGCCCCAGAGACCCAGCAGAGAAGAA 360
Db 301 TTCGACAGCTCCAGAGATCTTCCGAGATCATGCCCCAGAGACCCAGCAGAGAAGAA 360

QY 361 AATGTTTGAAGGAAGTGCAGAGCATCTGCCCCACAAAATGATGGGAAACAGCTGTGC 420
Db 361 AATGTTTGAAGGAAGTGCAGAGCATCTGCCCCACAAAATGATGGGAAACAGCTGTGC 420

QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCAAAGGGGG 480
Db 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCAAAGGGGG 480

QY 481 AAACATGCTGGACCCAGAGCTGCTGAGAGAGAGCTGCTGTTTATGAAGATC 540
Db 481 AAACATGCTGGACCCAGAGCTGCTGAGAGAGAGCTGCTGTTTATGAAGATC 540

QY 541 AGCGACCTTGAGGAAGATGACGAGTAATCCCTCTCG 576
Db 541 AGCGACCTTGAGGAAGATGACGAGTAATCCCTCTCG 576

RESULT 7
BD218420
LOCUS
DEFINITION BD218420
ACCESSION BD218420
VERSION JP 2002519013-A/5.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

BD218420 576 bp DNA linear PAT 17-JUL-2003
SSX gene, method of determining the occurrence of cancer in sample by determining the expression of peptides originating in the SSX gene and NY-ESO-1 gene and utilization thereof.

```

REFERENCE
AUTHORS Tureci,O., Sahin,U., Pfreundschuh,M., Rammensee,G., Stevanovic,S.,
Chen,Y.T., Gure,A. and Old,L.J.
TITLE SSX gene, method of determining the occurrence of cancer in sample
by determining the expression of peptides originating in the SSX
gene and NY-ESO-1 gene and utilization thereof
JOURNAL Patent: JP 2002519013-A 5 02-JUL-2002;
LUDWIG INSTITUTE FOR CANCER RESEARCH
COMMENT OS Homo sapiens (human)
PN JP 2002519013-A/5
PD 02-JUL-2002
PF 25-JUN-1999 JP 2000557145
PR 26-JUN-1998 US 09/105839
PI OZLEM TURECI,UGUR SAHIN,MICHAEL PFEUNDSCUH,GEORG RAMMENSEE,
STEFAN STEVANOVIC,YAO TSENG CHEN,ALI GURE,LLOYD J OLD PC
C12N15/09,A61K38/00,C07H21/04,C12N5/10,C12P21/04,C12Q1/58, PC
G01N33/48,
PC G01N33/53,G01N33/574,C12N15/00,A61K37/02,C12N5/00 CC SSX
gene, method of determining the occurrence of cancer in CC
sample by
CC determining the expression of peptides originating in the SSX
gene and
CC NY-ESO-1 gene and utilization thereof
FH Key Location/Qualifiers
FT source 1..576
FT /organism='Homo sapiens (human)'.
FEATURES
source 1..576
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAACCGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
DB 1 ATGAACCGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAAGCGCTTCGATGATATTGGCCAAATCTCTCTTAAGAAAGAGTGGGAAAAGATG 120
DB 61 TTACGAAAGCGCTTCGATGATATTGGCCAAATCTCTCTTAAGAAAGAGTGGGAAAAGATG 120
QY 121 AAATCCTCGGAGAAATCGTCTATGCTATATGAAGCTAACTATGAGGTCATGACTAAA 180
DB 121 AAATCCTCGGAGAAATCGTCTATGCTATATGAAGCTAACTATGAGGTCATGACTAAA 180
QY 181 CTAGGTTTCAAGGTCACCCCTCCCAAGCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGGTCACCCCTCCCAAGCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGGATGATTTGGTACGATCGAAGACACAGGAATCAGGTTGAAGTCTCTCAGATGACT 300
DB 241 GGGATGATTTGGTACGATCGAAGACACAGGAATCAGGTTGAAGTCTCTCAGATGACT 300
QY 301 TTGCGAGCTCCAGAGAAATCTCCCGAAGATCATGCCAGAGACCCAGAGAGAGAA 360
DB 301 TTGCGAGCTCCAGAGAAATCTCCCGAAGATCATGCCAGAGACCCAGAGAGAGAA 360
QY 361 AATGGTTTGAAGGAAGTGCAGAGGCTCTGGCCCAACAAATATGATGGAAACAGCTCTGC 420
DB 361 AATGGTTTGAAGGAAGTGCAGAGGCTCTGGCCCAACAAATATGATGGAAACAGCTCTGC 420
QY 421 CCCCCGGGAATCCAAAGTACCTTGGAGAAGATTAAAGACATCTGACCCCAAGGGGG 480
DB 421 CCCCCGGGAATCCAAAGTACCTTGGAGAAGATTAAAGACATCTGACCCCAAGGGGG 480
QY 481 AAACATGCTGGACCCACAGCTCGGTGAGAGAAGCAGCTGGTGTATGAGAGATC 540
DB 481 AAACATGCTGGACCCACAGCTCGGTGAGAGAAGCAGCTGGTGTATGAGAGATC 540
103 541 AGCGACCTCAGGAAGATGACGAGTACTCCCTCG 576
103 541 AGCGACCTCAGGAAGATGACGAGTACTCCCTCG 576
RESULT 8
HSU90841
LOCUS Homo sapiens SSX4 (SSX4) mRNA, complete cds. PRI 18-MAR-1998
DEFINITION Homo sapiens SSX4 (SSX4) mRNA, complete cds.
ACCESSION U90841
VERSION U90841.1 GI:2952022
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 576)
Gure,A.O., Tureci,O., Sahin,U., Tsang,S., Scanlan,M.J., Jager,E.,
Knuth,A., Pfreundschuh,M., Old,L.J. and Chen,Y.T.
SSX: a multigene family with several members transcribed in normal
testis and human cancer
Int. J. Cancer 72 (6), 965-971 (1997)
JOURNAL 98021352
MEDLINE 9378559
PUBMED
REFERENCE 2 (bases 1 to 576)
Gure,A.O., Tureci,O., Sahin,U., Tsang,S., Scanlan,M., Knuth,A.,
Pfreundschuh,M., Old,L.J. and Chen,Y.T.
Direct Submission
Submitted (05-FEB-1997) Pathology C-320, Cornell University, 1300
York Ave., New York City, NY 10021, USA
LOCATION/Qualifiers
1..576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..576
/genes="SSX4"
1..567
/genes="SSX4"
/notes="putative translation product; Kruppel-associated
box containing SSX gene"
/codon_start=1
/product="SSX4"
/protein_id="AAC05820.1"
/db_xref="GI:2952023"
/translation="MNGDDAFARPRDDAQISEKLKAFDDIAKYFSKWEKWKSE
KIVIVIMKLNIENVTKLGFVTLDPFMRSGRAADPHGNDGNDNRHNRQVERQMTEG
SQRIPIPKMPKPEEENGLKEVPBAGSGQNDGKQLCPGPNSTLEKINKTSGPKRG
KHAWTHLRERKQLVYVERISDPEDE"
1..18
/genes="SSX4"
/notes="corresponds to the 5' PCR primer that was used to
isolate this gene and is identical to SSX2, GenBank
Accession Number X86175; therefore, the SSX4 gene sequence
might contain differences within this span"
primer_bind
Query Match 100.0%; Score 576; DB 9; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAACCGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
DB 1 ATGAACCGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAAGCGCTTCGATGATATTGGCCAAATCTCTCTTAAGAAAGAGTGGGAAAAGATG 120
DB 61 TTACGAAAGCGCTTCGATGATATTGGCCAAATCTCTCTTAAGAAAGAGTGGGAAAAGATG 120
QY 121 AAATCCTCGGAGAAATCGTCTATGCTATGCTATGAGCTAACTATGAGGTCATGACTAAA 180
DB 121 AAATCCTCGGAGAAATCGTCTATGCTATGCTATGAGCTAACTATGAGGTCATGACTAAA 180

```

```
181 CTAGGTTTCAAGGTCACACCTCCACCTTTCATGGTAGTAAACGGGCTGCAGACTTCCAC 240
Db
181 CTAGGTTTCAAGGTCACACCTCCACCTTTCATGGTAGTAAACGGGCTGCAGACTTCCAC 240
Qy
241 GGGAAATGATTTGGTAGAGTACGAAACACACAGGAATCAGGTTGAAGCTCTCAGATGACT 300
Db
241 GGGAAATGATTTGGTAGAGTACGAAACACACAGGAATCAGGTTGAAGCTCTCAGATGACT 300
Qy
301 TTCCGGCAGCCCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAAGCCAGCAGAGGAAGAA 360
Db
301 TTCCGGCAGCCCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAAGCCAGCAGAGGAAGAA 360
Qy
361 AATGTTTGAAGAGTCCAGAGCATCTGCGCCACAAATATGAGGAACACGCTGTGC 420
Db
361 AATGTTTGAAGAGTCCAGAGCATCTGCGCCACAAATATGAGGAACACGCTGTGC 420
Qy
421 CCCCCGGGAATCAAGTACCTTGGAGAGATTACAGACATCTGGACCCAAAAGGGGG 480
Db
421 CCCCCGGGAATCAAGTACCTTGGAGAGATTACAGACATCTGGACCCAAAAGGGGG 480
Qy
481 AAACATGCTCGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
Db
481 AAACATGCTCGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
Qy
541 AGCAGCCTCGAGGAAGATGACGAGTAATCTCCCTCG 576
Db
541 AGCAGCCTCGAGGAAGATGACGAGTAATCTCCCTCG 576
```

RESULT 9

```
BC005325 1250 bp mRNA linear PRI 03-OCT-2003
LOCUS Homo sapiens synovial sarcoma, X breakpoint 4, transcript variant
DEFINITION 1. mRNA (cdna clone MGC:12411 IMAGE:3961688), complete cds.
```

BC005325

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1250)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Schneitz, R.E., Brownstein, M.J., Usdin, T.B., Ioshizuka, S.,

Carninci, P., Frange, C., Raha, S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smallos, D.E.,

Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 1250)

Strausberg, R.

Direct Submission

Submitted (27-MAR-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

REMARK

COMMENT

Contact: MGC help desk
Email: cgabs-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAL Plate: 16 Row: m Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 28559011.

FEATURES

source

1. .1250
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12411 IMAGE:3961688"
/tissue_type="Bladder, carcinoma"
/clone_lib="NIH_MGC_53"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
1. .1250
/gene="SSX4"
/note="synonym: MGC12411"
/db_xref="LocusID:6759"
/db_xref="MIM:300326"
59. .625
/codon_start=1
/product="synovial sarcoma, X breakpoint 4, isoform a"
/protein_id="AAH05325.1"
/db_xref="GI:13529095"
/translation="MNGDDAFARRPRDDAQISEKLKAFDDIAKYFKSEWKMSSE
KIVVYMKLVNVEVTKLGPVTLPPFMSKRAADFGNDFGNDRNHRNQVERPOMTFG
SLQRIPIKMPKPEASEENGLKEVPEASGPQNDGKQLCPFPNPSTLEKINKTSFKRG
KHAWTHLRERKQLVVEEISDFEEDDE"
125. .304
/misc_feature
/note="KRAB; Region: krueppel associated box"
/db_xref="CDD:smart00349"

CDS

Query Match 99.7%; Score 574.4; DB 9; Length 1250;
Best Local Similarity 99.8%; Pred. No. 8.8e-164;
Matches 575; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAACGAGACGACGCGCTTTGCAAGGAGAGACCCAGGAGATGCTCAAAATATCAGAGAAG 60
Db 59 ATGAACGAGACGACGCGCTTTGCAAGGAGAGACCCAGGAGATGCTCAAAATATCAGAGAAG 118
Qy 61 TTACGAAAGCGCTTCGATGATATTCGCAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 119 TTACGAAAGCGCTTCGATGATATTCGCAATATCTTCTTAAGAAAGAGTGGGAAAGATG 178
Qy 121 AAATCCTCGGAGAAATCGTCTATGTATGATGAAGCTAAACTATGAGGTTCATGACTAAA 180
Db 179 AAATCCTCGGAGAAATCGTCTATGTATGATGAAGCTAAACTATGAGGTTCATGACTAAA 238
Qy 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 239 CTAGGTTTCAAGGTCACCTCCACCTTTCATGGTAGTAAACGGGCTGCAGACTTCCAC 298
Qy 241 GGGAAATGATTTGGTAGAGTACGAAACACACAGGAATCAGGTTGAAGCTCTCAGATGACT 300
Db 299 GGGAAATGATTTGGTAGAGTACGAAACACACAGGAATCAGGTTGAAGCTCTCAGATGACT 358
Qy 301 TTCCGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAAGCCAGCAGAGGAAGAA 360

ORIGIN

Query Match 99.7%; Score 574.4; DB 9; Length 1250;
Best Local Similarity 99.8%; Pred. No. 8.8e-164;
Matches 575; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAACGAGACGACGCGCTTTGCAAGGAGAGACCCAGGAGATGCTCAAAATATCAGAGAAG 60
Db 59 ATGAACGAGACGACGCGCTTTGCAAGGAGAGACCCAGGAGATGCTCAAAATATCAGAGAAG 118
Qy 61 TTACGAAAGCGCTTCGATGATATTCGCAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 119 TTACGAAAGCGCTTCGATGATATTCGCAATATCTTCTTAAGAAAGAGTGGGAAAGATG 178
Qy 121 AAATCCTCGGAGAAATCGTCTATGTATGATGAAGCTAAACTATGAGGTTCATGACTAAA 180
Db 179 AAATCCTCGGAGAAATCGTCTATGTATGATGAAGCTAAACTATGAGGTTCATGACTAAA 238
Qy 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 239 CTAGGTTTCAAGGTCACCTCCACCTTTCATGGTAGTAAACGGGCTGCAGACTTCCAC 298
Qy 241 GGGAAATGATTTGGTAGAGTACGAAACACACAGGAATCAGGTTGAAGCTCTCAGATGACT 300
Db 299 GGGAAATGATTTGGTAGAGTACGAAACACACAGGAATCAGGTTGAAGCTCTCAGATGACT 358
Qy 301 TTCCGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAAGCCAGCAGAGGAAGAA 360

```
Db 359 TTGGGAGCCTCCAGAGAAATCTCCGGAAGATCATGCCAAGAGCCAGCAGAGGAAGAA 418
QY 361 AATGGTTTGAAGGAAGTGCAGAGGCATCTGGCCCAACAAATATGGAACACACTGTGC 420
Db 419 AATGGTTTGAAGGAAGTGCAGAGGCATCTGGCCCAACAAATATGGAACACACTGTGC 478
QY 421 CCCCCGGGAAATCCAAGTACCTTGGAGAAGATTAAAGACATCTGGACCCCAAAAGGGG 480
Db 479 CCCCCGGGAAATCCAAGTACCTTGGAGAAGATTAAAGACATCTGGACCCCAAAAGGGG 538
QY 481 AAACATGCTGGAGCCCAAGACATCTGGAGAGAGAGAGAGAGAGAGAGAGAGATC 540
Db 539 AAACATGCTGGAGCCCAAGACATCTGGAGAGAGAGAGAGAGAGAGAGAGATC 598
QY 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 576
Db 599 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 634

RESULT 10
LOCUS A48452 766 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 17 from Patent WO9602641.
ACCESSION A48452
VERSION A48452.1 GI:2302236
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 766)
AUTHORS Cooper, C.S. and Gusterson, B.A.
TITLE MATERIALS AND METHODS RELATING TO THE DIAGNOSIS AND PROPHYLACTIC
AND THERAPEUTIC TREATMENT OF SYNOVIAL SARCOMA
JOURNAL Patent: WO 9602641-A 17 01-FEB-1996;
CANCER RES CAMPAIGN TECH (GB)
COMMENT Other publication AU 2986595 960216.
FEATURES
Location/Qualifiers
source 1..766
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 86.2%; Score 496.6; DB 6; Length 766;
Best Local Similarity 91.5%; Pred. No. 4.6e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTGCAAGAGAGACCCAGGATGATGCTCAAAATACAGAGAAG 60
Db 92 ATGAACGGAGACGACGCTTTGCAAGAGAGACCCAGGTTGGTGTCTCAAAATACAGAGAAG 151
QY 61 TTACGAAAGGCTTCGATGATTTGCCAAATATCTCTTAAGAAAGAGTGGAAAGATG 120
Db 152 ATCCAAAAGGCTTCGATGATTTGCCAAATATCTCTTAAGAAAGAGTGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGGTCTAGACTAAA 180
Db 212 AAAGCCTCGAGAAATCTCTATGTATATGAAGAAAGATGAGGCTATGACTAAA 271
QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTCATGTTAGTAAACGGGTGAGACTTCCAC 240
Db 272 CTAGGTTTCAAGGCAACCTCCACCTTTCATGTTAGTAAACGGGTGAGACTTCCAG 331
QY 241 GGGAAATGATTTGGTAACGATCGAAACCAAGGAATCAGGTTGAACGCTCTCAGACT 300
Db 332 GGGAAATGATTTGGTAACGATCGAAACCAAGGAATCAGGTTGAACGCTCTCAGACT 391
QY 301 GGGAAATGATTTGGTAACGATCGAAACCAAGGAATCAGGTTGAACGCTCTCAGACT 360
Db 392 GGGAAATGATTTGGTAACGATCGAAACCAAGGAATCAGGTTGAACGCTCTCAGACT 451
QY 301 TTGGGAGCCTCCAGAGAAATCTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
Db 392 TTGGGAGCCTCCAGAGAAATCTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 451
QY 361 AATGGTTTGAAGGAAGTGCAGAGGCATCTGGCCCAACAAATATGGAACACACTGTGC 420
Db 419 AATGGTTTGAAGGAAGTGCAGAGGCATCTGGCCCAACAAATATGGAACACACTGTGC 478
QY 421 CCCCCGGGAAATCCAAGTACCTTGGAGAAGATTAAAGACATCTGGACCCCAAAAGGGG 480
Db 479 CCCCCGGGAAATCCAAGTACCTTGGAGAAGATTAAAGACATCTGGACCCCAAAAGGGG 538
QY 481 AAACATGCTGGAGCCCAAGACATCTGGAGAGAGAGAGAGAGAGAGAGAGATC 540
Db 539 AAACATGCTGGAGCCCAAGACATCTGGAGAGAGAGAGAGAGAGAGAGAGATC 598
QY 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 576
Db 599 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 634
```

```
Db 452 AATGATTCGGAGCAAGTGCAGAAAGCATCTGGCCCAACAAATATGATGGAAAGAGCTGTGC 511
QY 421 CCCCCGGGAAATCCAAGTACCTTGGAGAAGATTAAAGACATCTGGACCCCAAAAGGGG 480
Db 512 CCCCCGGGAAACCAACTCTCTGAGAAGATTACAGAGATCTGGACCCCAAAAGGGG 571
QY 481 AAACATGCTGGAGCCCAAGACATCTGGAGAGAGAGAGAGAGAGAGAGATC 540
Db 572 GAACATGCTGGAGCCCAAGACATCTGGAGAGAGAGAGAGAGAGAGATC 631
QY 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 575
Db 632 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 666

RESULT 11
LOCUS AX821947 766 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 75 from Patent WO03068961.
ACCESSION AX821947
VERSION AX821947.1 GI:39725168
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Andrews, P.A., Waleh, J.A. and Gokhale, P.A.
TITLE Method to modify differentiation of pluripotential stem cells
JOURNAL Patent: WO 03068961-A 75 21-AUG-2003;
Axordia Limited (GB)
FEATURES
Location/Qualifiers
source 1..766
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 86.2%; Score 496.6; DB 6; Length 766;
Best Local Similarity 91.5%; Pred. No. 4.6e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTGCAAGAGAGACCCAGGATGATGCTCAAAATACAGAGAAG 60
Db 92 ATGAACGGAGACGACGCTTTGCAAGAGAGACCCAGGTTGGTGTCTCAAAATACAGAGAAG 151
QY 61 TTACGAAAGGCTTCGATGATTTGCCAAATATCTCTTAAGAAAGAGTGGAAAGATG 120
Db 152 ATCCAAAAGGCTTCGATGATTTGCCAAATATCTCTTAAGAAAGAGTGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGGTCTAGACTAAA 180
Db 212 AAAGCCTCGAGAAATCTCTATGTATATGAAGAAAGATGAGGCTATGACTAAA 271
QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTCATGTTAGTAAACGGGTGAGACTTCCAC 240
Db 272 CTAGGTTTCAAGGCAACCTCCACCTTTCATGTTAGTAAACGGGTGAGACTTCCAG 331
QY 241 GGGAAATGATTTGGTAACGATCGAAACCAAGGAATCAGGTTGAACGCTCTCAGACT 300
Db 332 GGGAAATGATTTGGTAACGATCGAAACCAAGGAATCAGGTTGAACGCTCTCAGACT 391
QY 301 TTGGGAGCCTCCAGAGAAATCTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
Db 392 TTGGGAGCCTCCAGAGAAATCTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 451
QY 361 AATGGTTTGAAGGAAGTGCAGAGGCATCTGGCCCAACAAATATGGAACACACTGTGC 420
Db 452 AATGATTCGGAGGAAGTGCAGAGGCATCTGGCCCAACAAATATGGAACACACTGTGC 511
QY 421 CCCCCGGGAAATCCAAGTACCTTGGAGAGAGATTAAAGACATCTGGACCCCAAAAGGGG 480
Db 512 CCCCCGGGAAACCAACTCTCTGAGAAGATTACAGAGAGATCTGGACCCCAAAAGGGG 571
```

```
QY 481 AAACATGCTCGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
|
|
|
Db 572 GAACATGCTCGGACCCACAGACTGCGTGAGAGAAACAGCTGGTATGAAGAGATC 631
|
|
|
QY 541 AGCGACCTGAGGAAGATGACGAGTAACTCCCTC 575
|
|
|
Db 632 AGCGACCTGAGGAAGATGACGAGTAACTCCCTC 666
|
|
|
RESULT 12
HSSX2
LOCUS HSSX2 766 bp mRNA linear PRI 02-FEB-1999
DEFINITION H.sapiens mRNA for SSX2 protein.
ACCESSION X86175
VERSION X86175.1 GI:829113
KEYWORDS SSX1 gene; SSX2 gene; synovial sarcoma.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Crew A.J., Clark J., Fisher C., Gill S., Grimer R., Chand A., Shipley J., Gusterson B.A. and Cooper C.S.
TITLE Fusion of SYR to two genes, SSX1 and SSX2, encoding proteins with homology to the Kruppel-associated box in human synovial sarcoma
JOURNAL EMBO J. 14 (10), 2333-2340 (1995)
MEDLINE 95292974
PUBMED 7539744
REFERENCE 2 (bases 1 to 766)
AUTHORS Cooper, C.S.
TITLE Direct Submission
JOURNAL Submitted (10-APR-1995) C.S. Cooper, Institute of Cancer Research, Haddow Laboratories, Cotswold Road, Sutton, Surrey SM2 5NG, UK
FEATURES
source
1..766
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="adult"
2..766
/gene="SSX2"
20..61
/gene="SSX2"
/misc_feature
92..658
/note="Kruppel associated box homology"
CDS
92..658
/gene="SSX2"
/codon_start=1
/protein_id="CAA60111.1"
/db_xref="GI:4218445"
/db_xref="GOA:Q16385"
/translacion="MNGDAPARPTVGAQIPEIKIQAFFDIKAFYKSEWEKMKASE
KIFVYMKRYEAMTKLGPATLPFMCNKAEDFGNDLNDNRGNQVERPOMTFG
RLQISPKMPKXPAESGNDSEEVPEASGPQNDKELCPGKPTTSEKIHRSFPKG
EHWATHRLERKQLVITEISDPEDDE"
ORIGIN
Query Match 86.2%; Score 496.6; DB 9; Length 766;
Best Local Similarity 91.5%; Pred. No. 4.6e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGCAGCGCTTTGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 92 ATGAACGGAGCAGCGCTTTGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151
QY 61 TTACGAAGAGCGCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGAGAGAGAGAGAGAG 120
Db 152 ATCCAAAAGCGCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGAGAGAGAGAGAGAG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTGTATGAAGCTAAACTATGAGGTCTATGACTAAA 180
Db 212 AAAGCTCGGAGAAATCTCTATGTGTATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 271
```

```
QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTCAATGCGTAGTAAACGGCTCGAGACTTCCAC 240
|
|
|
Db 272 CTAGGTTTCAAGGCGACCCCTCCACCTTTCAATGCGTAGTAAACGGCTCGAGACTTCCAG 331
|
|
|
QY 241 GGAATGATTTGGTTAAAGATGAACACACAGGATCAGTTTGAAGCTCTCAGATGACT 300
|
|
|
Db 332 GGAATGATTTGGTTAAAGATGAACCTAACCGTGGGATCAGTTTGAAGCTCTCAGATGACT 391
|
|
|
QY 301 TTCGGCAGCTCCAGAGAATCTTCCGAAGATCATGCCAAGAGCCAGCAGAGGAAGAA 360
|
|
|
Db 392 TTCGGCAGCTCCAGGGAATCTCCCGAGAGATCATGCCAAGAGCCAGCAGAGGAAGAA 451
|
|
|
QY 361 AATGTTTGAAGAAATGCGCAGAGGATCTGGCCCAAAATGATGGGAAAAGCTGTGC 420
|
|
|
Db 452 AATGATTCGAGGAAGTGCAGAAAGCATCTGGCCCAAAATGATGGGAAAAGCTGTGC 511
|
|
|
QY 421 CCCCCGGAAATCCAAAGTACCTTGGAGAAGATTAAACAAGACATCTCGACCCAAAGGGG 480
|
|
|
Db 512 CCCCCGGAAATCCAAAGTACCTTGGAGAAGATTAAACAAGACATCTCGACCCAAAGGGG 571
|
|
|
QY 481 AAACATGCTCGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
|
|
|
Db 572 GAACATGCTCGGACCCACAGACTGCGTGAGAGAAACAGCTGGTGGTTTATGAAGAGATC 631
|
|
|
QY 541 AGCGACCTGAGGAAGATGACGAGTAACTCCCTC 575
|
|
|
Db 632 AGCGACCTGAGGAAGATGACGAGTAACTCCCTC 666
|
|
|
RESULT 13
AR025465
LOCUS AR025465 931 bp DNA linear PAT 05-DEC-1998
DEFINITION Sequence 2 from patent US 5798264.
ACCESSION AR025465
VERSION AR025465.1 GI:3978093
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 931)
AUTHORS Pfreundschuh M.
TITLE Isolated nucleic acid molecules which encode renal cancer specific antigens, and uses thereof
JOURNAL Patent: US 5798264-A 2 25-AUG-1998;
FEATURES Location/Qualifiers
source
1..931
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 86.2%; Score 496.6; DB 6; Length 931;
Best Local Similarity 91.5%; Pred. No. 4.7e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGCAGCGCTTTGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 99 ATGAACGGAGCAGCGCTTTGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 158
QY 61 TTACGAAGAGCGCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGAGAGAGAGAGAGAG 120
Db 159 ATCCAAAAGCGCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGAGAGAGAGAGAGAG 218
QY 121 AAATCCTCGGAGAAATCGTCTATGTGTATGAAGCTAAACTATGAGGTCTATGACTAAA 180
Db 219 AAAGCTCGGAGAAATCTCTATGTGTATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 278
QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTCAATGCGTAGTAAACGGCTCGAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGCGACCCCTCCACCTTTCAATGCGTAGTAAACGGCTCGAGACTTCCAG 338
QY 241 GGAATGATTTGGTTAAAGATGAACCTAACCGTGGGATCAGTTTGAAGCTCTCAGATGACT 300
|
|
|
```



```
Db 339 GGGATGATTGGATAATGACCCCTAACCGTGGGAATCAGGTTGAAAGCTCCTCAGATGACT 398
QY 301 TTCGGAGGCTCCAGAGATCTTCCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
Db 399 TTCGGAGGCTCCAGGGAATCTCCCGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 458
QY 361 AATGGTTTGAAGGAAGTGCAGAGGATCTTGCCCAACAAATGATGGGAAACAGCTGTGC 420
Db 459 AATGATTCGGAGGAAGTGCAGAGGATCTTGCCCAACAAATGATGGGAAAGAGCTGTGC 518
QY 421 CCCCGGGAATCCAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 480
Db 519 CCCCGGGAATCCAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 578
QY 481 AAACATGCTCGGAGCCACACAGCTGCTGAGAGAAAGCAGCTGCTGATTAAGAGATC 540
Db 579 GAACATGCTCGGAGCCACACAGCTGCTGAGAGAAAGCAGCTGCTGATTAAGAGATC 638
QY 541 AGGACCCCTGAGGAAGATGACGAGTAACCTCCCTC 575
Db 639 AGGACCCCTGAGGAAGATGACGAGTAACCTCCCTC 673

RESULT 15
LOCUS AR117874 931 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5840568.
ACCESSION AR117874
VERSION AR117874.1 GI:5986830
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 931)
AUTHORS Pfreundschuh,M.
TITLE Hodgkin's disease associated molecules and uses thereof
JOURNAL Patent: US 5840568-A 2 24-NOV-1998;
FEATURES
    Location/Qualifiers
    source
    1..931
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN

Query Match 86.2%; Score 496.6; DB 6; Length 931;
Best Local Similarity 91.5%; Pred. No. 4.7e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTGGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 99 ATGAACGGAGACGACGCTTGGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGCCCTTCGATGATATTGCCAAATACTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 159 ATCCAAAGCCCTTCGATGATATTGCCAAATACTTCTTAAGAAAGAGTGGGAAAGATG 218
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCATGACTAAA 180
Db 219 AAAGCCTCGGAGAAATCTTCTATGTATATGAAGAAAGATGAGGCTATGACTAAA 278
QY 181 CTAGGTTTCAAGGTCACCCCTCCACCTTTCATCGTAGTAAACGGGTCGAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGTCACCCCTCCACCTTTCATCGTAGTAAACGGGTCGAGACTTCCAG 338
QY 241 GGAATGATTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGTCCTCAGTACT 300
Db 339 GGAATGATTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGTCCTCAGTACT 398
QY 301 TTCGGAGGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
Db 399 TTCGGAGGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 458
QY 361 AATGGTTTGAAGGAAGTGCAGAGGATCTTGCCCAACAAATGATGGGAAACAGCTGTGC 420
Db 459 AATGATTCGGAGGAAGTGCAGAGGATCTTGCCCAACAAATGATGGGAAAGAGCTGTGC 518
QY 421 CCCCGGGAATCCAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 480
Db 519 CCCCGGGAATCCAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 578
QY 361 AATGGTTTGAAGGAAGTGCAGAGGATCTTGCCCAACAAATGATGGGAAACAGCTGTGC 420
```

```
Db 459 AATGATTCGGAGGAAGTGCAGAGAATCTGGCCCAACAAATGATGGGAAAGAGCTGTGC 518
QY 421 CCCCGGGAATCCAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 480
Db 519 CCCCGGGAATCCAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 578
QY 481 AAACATGCTCGGAGCCACACAGCTGCTGAGAGAAAGCAGCTGCTGATTAAGAGATC 540
Db 579 GAACATGCTCGGAGCCACACAGCTGCTGAGAGAAAGCAGCTGCTGATTAAGAGATC 638
QY 541 AGGACCCCTGAGGAAGATGACGAGTAACCTCCCTC 575
Db 639 AGGACCCCTGAGGAAGATGACGAGTAACCTCCCTC 673

RESULT 15
LOCUS AR117874 931 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 2 from patent US 6140464.
ACCESSION AR117874
VERSION AR117874.1 GI:14098780
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 931)
AUTHORS Pfreundschuh,M. and Rammensee,H.-G.
TITLE Nonapeptides that bind a HLA-A2.1 molecule
JOURNAL Patent: US 6140464-A 2 31-OCT-2000;
FEATURES
    Location/Qualifiers
    source
    1..931
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN

Query Match 86.2%; Score 496.6; DB 6; Length 931;
Best Local Similarity 91.5%; Pred. No. 4.7e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTGGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 99 ATGAACGGAGACGACGCTTGGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGCCCTTCGATGATATTGCCAAATACTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 159 ATCCAAAGCCCTTCGATGATATTGCCAAATACTTCTTAAGAAAGAGTGGGAAAGATG 218
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCATGACTAAA 180
Db 219 AAAGCCTCGGAGAAATCTTCTATGTATATGAAGAAAGATGAGGCTATGACTAAA 278
QY 181 CTAGGTTTCAAGGTCACCCCTCCACCTTTCATCGTAGTAAACGGGTCGAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGTCACCCCTCCACCTTTCATCGTAGTAAACGGGTCGAGACTTCCAG 338
QY 241 GGAATGATTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGTCCTCAGTACT 300
Db 339 GGAATGATTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGTCCTCAGTACT 398
QY 301 TTCGGAGGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
Db 399 TTCGGAGGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 458
QY 361 AATGGTTTGAAGGAAGTGCAGAGGATCTTGCCCAACAAATGATGGGAAACAGCTGTGC 420
Db 459 AATGATTCGGAGGAAGTGCAGAGGATCTTGCCCAACAAATGATGGGAAAGAGCTGTGC 518
QY 421 CCCCGGGAATCCAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 480
Db 519 CCCCGGGAATCCAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 578
QY 481 AAACATGCTCGGAGCCACACAGCTGCTGAGAGAAAGCAGCTGCTGATTAAGAGATC 540
```

Db 579 GAACATGCTGGACCCACAGACTGCGTGAGAGAAAAACAGCTGCTGCTGATTATGAAGAGATC 638
QY 541 AGCGACCCCTGAGGAAGATGACGAGTAACTCCCTC 575
Db 639 AGCGACCCCTGAGGAAGATGACGAGTAACTCCCTC 673

Search completed: March 31, 2004, 16:39:15
Job time : 3494 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 13:52:29 ; Search time 437 Seconds
(without alignments)
5599.456 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 576

Sequence: 1 ATGAACGGACGACGACGCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	576	100.0	576	6	ABK84472 Human cDN
2	576	100.0	576	7	ABQ83858 Human SSX
3	576	100.0	1250	9	ADC09598 SSX-4 cDN
4	574.4	99.7	576	2	Aav70060 Human SSX
5	496.6	86.2	766	6	Aat11780 Human X-c
6	496.6	86.2	766	6	ABs73282 DNA encod
7	496.6	86.2	766	7	ABs73282 Human col
8	496.6	86.2	766	7	ABQ83844 Human SSX
9	496.6	86.2	766	9	ADC09570 SSX-2 cDN
10	496.6	86.2	766	9	ADD25523 Binding d
11	496.6	86.2	931	2	Aat47748 Melanoma
12	496.6	86.2	931	2	Aav04267 Melanoma
13	496.6	86.2	1309	6	ABL64122 Breast ca
14	496.6	86.2	1309	6	ABL63730 Breast ca
15	487.2	84.6	921	5	AA80402 DNA encod
16	487.2	84.6	921	5	AA80399 DNA encod
17	480.6	83.4	576	2	AAV70061 Human SSX
18	480	83.3	766	2	Aat11779 Human X-c
19	221.4	38.4	711	2	Aat11781 Human SYT
20	203.6	36.4	585	2	Aat11782 Human SYT
21	192.6	33.4	830	5	AA80401 DNA encod
22	191.4	33.2	822	5	AA80404 DNA encod
23	117.4	20.4	165	6	ABs73281 DNA encod

24	74.2	12.9	1110	5	AAS91978	Aas91978 DNA encod
25	42.2	7.3	2000	7	ADA71938	Ada71938 Rice Gene
26	41.4	7.2	1981	9	ADC32171	Adc32171 Human nov
27	40.8	7.1	80	6	ABN33980	Abn33980 Human spl
c 28	40.8	7.1	8059	3	AAA81747	Aaa81747 N. mening
29	40.8	7.1	110000	3	AAA81490	Continuation (8 of
30	40.8	7.1	349980	3	AAF21608	AAF21608 Neisseria
31	39.8	6.9	571	6	ABK71614	Abk71614 Human dit
32	39.8	6.9	3591	8	ADA10966	Ada10966 Human cDN
c 33	39.8	6.9	3591	8	ACH04006	ACH04006 Human cDN
34	39.4	6.8	1097	4	AAF27660	AAF27660 DNA encod
35	39.4	6.8	1097	8	ACH04007	ACH04007 Human cDN
36	37.8	6.6	2736	4	AAI57976	AAI57976 Human pol
37	37.8	6.6	2872	4	AAH98654	Aah98654 Human EST
38	37.8	6.6	2872	4	AAI59762	Aai59762 Human pol
39	37.8	6.6	5532	9	ADB47398	Adb47398 Human cDN
40	37.4	6.5	730	4	AAC91340	Aac91340 Human pol
41	37	6.4	653	6	ABQ99153	Abq99153 Human ORF
42	36	6.2	2369	6	ABA01156	Abao1156 Human zin
43	36	6.2	2798	9	ADC30277	Adc30277 Human nov
c 44	36	6.2	349980	5	AAH41224	Aah41224 Pyrococcu
45	35.8	6.2	765	4	AAK91889	Aak91889 Human cDN

ALIGNMENTS

RESULT 1
ABK84472
ID ABK84472 standard; cDNA; 576 BP.
XX AC ABK84472;
DT 14-AUG-2002 (first entry)
XX DE Human cDNA differentially expressed in granulocytic cells #1043.
XX KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX OS Homo sapiens.
XX WC200228999-A2.
XX 11-APR-2002.
XX PD 03-OCT-2001; 2001WO-US030821.
XX PF 03-OCT-2000; 2000US-0237189P.
XX PR (GENE-) GENE LOGIC INC.
XX PA Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
PI WPI; 2002-435328/46.
XX DR Detecting granulocyte activation by detecting differential expression of
XX genes associated with granulocyte activation, which serves as diagnostic
XX markers that is useful for monitoring disease states and drug toxicity.
XX Claim 1; SEQ ID NO 1043; 114pp; English.
XX The invention relates to detecting (M1) granulocyte (GC) activation
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by
XX DNA chip analysis as given in the specification, and comparing the
XX expression level to an expression level in an unactivated GC, where
XX differential expression of Gs is indicative of GCA. Also included are

modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 576 BP; 187 A; 127 C; 150 G; 112 T; 0 U; 0 Other;

Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 9.3e-172;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGAGCGCCCTTTGCAAGGAGAGCCAGGGATGATGCTCAAAATATCAGAGAAG 60
Db 1 ATGAACGGAGAGCGCCCTTTGCAAGGAGAGCCAGGGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAGGCTTCGATGATATCCCAATCTCTCTAGAAAGAGTGGGAAAGATG 120
Db 61 TTACGAAGGCTTCGATGATATCCCAATCTCTCTAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAAATCTAGGTCATGACTAAA 180
Db 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAAATCTAGGTCATGACTAAA 180
QY 181 CTAGTTTCAAGGTCAACCTCCACCTTTCTATGCTAGTAAACCGGGCTGCAGACTTCCAC 240
Db 181 CTAGTTTCAAGGTCAACCTCCACCTTTCTATGCTAGTAAACCGGGCTGCAGACTTCCAC 240
QY 241 GGGATGATTTGGTAAAGTCGAAACACAGGATCAGGTTGACGTCCTCAGATGACT 300
Db 241 GGGATGATTTGGTAAAGTCGAAACACAGGATCAGGTTGACGTCCTCAGATGACT 300
QY 301 TTGGCAGCTCCAGAGATCTTCCGGAAGATCATGCCCAAGACCCAGCAGAGAGAA 360
Db 301 TTGGCAGCTCCAGAGATCTTCCGGAAGATCATGCCCAAGACCCAGCAGAGAGAA 360
QY 361 AATGTTTGAAGAGAGTCCAGAGGATCTTGSCCCACAAATATGATGGGAAACAGCTGTGC 420
Db 361 AATGTTTGAAGAGAGTCCAGAGGATCTTGSCCCACAAATATGATGGGAAACAGCTGTGC 420
QY 421 CCCCCTGGGAAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 480
Db 421 CCCCCTGGGAAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 480
QY 481 AAACATGCTGACCCACAGATCTGGTGAAGAGAGAGAGCTGGTGGTTTATGAAGATC 540
Db 481 AAACATGCTGACCCACAGATCTGGTGAAGAGAGAGAGCTGGTGGTTTATGAAGATC 540

QY 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576
Db 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576

RESULT 2

ABQ83858
ID ABQ83858 standard; cDNA; 576 BP.
XX ABQ83858;
AC ABQ83858;
XX 03-FEB-2003 (first entry)
DT Human SSX-4 encoding cDNA SEQ ID NO:599.
DE Human; epitope; vaccine; immunotherapeutic; cytostatic; immunogenicity;
KW T cell; gene; ss.
XX Homo sapiens.
OS WO200281646-A2.
PN 17-OCT-2002.
PD 04-APR-2002; 2002WO-US011101.
PF 06-APR-2001; 2001US-0282211P.
PR 07-NOV-2001; 2001US-0337017P.
PR 07-MAR-2002; 2002US-0363210P.
XX (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX Simard JUL, Diamond DC, Liu L, Xie Z;
XX WPI; 2003-067518/06.
XX P-PSDB; ABP74710.
PT Novel epitopes useful as vaccines, comprises peptides or nucleic acid
PT encoding the peptides, that are useful epitopes of target-associated
PT antigens.
XX Claim 1; Page 185; 352pp; English.
PS The present invention describes an isolated epitope (I) and an epitope
CC cluster. Also described is a vaccine or immunotherapeutic composition
CC (VC) comprising (I). (I) has cytostatic activity. VC is useful for
CC treating an animal, by administering to an animal the vaccine or
CC immunotherapeutic composition. VC is also useful for evaluating
CC immunogenicity of a vaccine or immunotherapeutic composition, by
CC administering VC to an HLA-transgenic animal and evaluating
CC immunogenicity based on a characteristic of the animal, or by in vitro
CC primary stimulation of a T cell and evaluating immunogenicity. (I) is
CC useful for determining specific T cell frequency, by contacting T cells
CC with a MHC-peptide complex, and further comprises ELISPOT analysis,
CC limiting dilution analysis, flow cytometry, in situ hybridisation and/or
CC polymerase chain reaction (PCR). ABQ83843 to ABQ83858 and ABP74128 to
CC ABP74713 represent sequences used in the exemplification of the present
CC invention
XX Sequence 576 BP; 187 A; 127 C; 150 G; 112 T; 0 U; 0 Other;

Query Match 100.0%; Score 576; DB 7; Length 576;
Best Local Similarity 100.0%; Pred. No. 9.3e-172;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGAGCGCTTTGCAAGGAGAGCCAGGGATGATGCTCAAAATATCAGAGAAG 60
Db 1 ATGAACGGAGAGCGCTTTGCAAGGAGAGCCAGGGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAGGCTTCGATGATATTCGCAAAATCTCTCTAGAAAGAGTGGGAAAGATG 120
Db 61 TTACGAAGGCTTCGATGATATTCGCAAAATCTCTCTAGAAAGAGTGGGAAAGATG 120

```

QY 121 AAATCTCGGAGAAATCGTATGTGTATATGAAGCTAAACTATGAGGTCAATGACTAAA 180
Db 121 AAATCTCGGAGAAATCGTATGTGTATATGAAGCTAAACTATGAGGTCAATGACTAAA 180
QY 181 CTAGGTTTCAGGTCACGCTCCCACTTTCATGCGTAGTAAACGGGCTGAGACTTCCAC 240
Db 181 CTAGGTTTCAGGTCACGCTCCCACTTTCATGCGTAGTAAACGGGCTGAGACTTCCAC 240
QY 241 GGGAAATGATTTGGTACGATCGAAACCCACAGGAATCAGGTTGAACGTCTCTCAGTACT 300
Db 241 GGGAAATGATTTGGTACGATCGAAACCCACAGGAATCAGGTTGAACGTCTCTCAGTACT 300
QY 301 TTCGGCAGCCTCCAGAGAAATCTTCCGAGAGATCATGCCAAGATCATGCCAAGAGCCAGCAGAGAA 360
Db 301 TTCGGCAGCCTCCAGAGAAATCTTCCGAGAGATCATGCCAAGATCATGCCAAGAGCCAGCAGAGAA 360
QY 361 AATGGTTTGAAGGAAGTCCAGAGGATCTGGCCCAAAATGATGGAAACAGCTGTGC 420
Db 361 AATGGTTTGAAGGAAGTCCAGAGGATCTGGCCCAAAATGATGGAAACAGCTGTGC 420
QY 421 CCCCCGGGAAATCCAACTACCTTGGAGAAGATTAAACAAGACATCTGGACCCCAAAAGGGG 480
Db 421 CCCCCGGGAAATCCAACTACCTTGGAGAAGATTAAACAAGACATCTGGACCCCAAAAGGGG 480
QY 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
Db 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
QY 541 AGCGACCTCGAGGAAGATGAGGTAATCTCCCTCG 576
Db 541 AGCGACCTCGAGGAAGATGAGGTAATCTCCCTCG 576

```

RESULT 3

ADCO9598
ID ADC09598 standard; cDNA; 1250 BP.

XX AC ADC09598;

XX DT 18-DEC-2003 (first entry)

XX DE SSX-4 cDNA #SEQ ID 599.

XX KW Epitope; immunological; vaccine;
XX KW major histocompatibility complex class I; MHC class I; cancer;
XX KW immunisation; ss.

XX OS Unidentified.

XX FN WO2003008537-A2.

XX PD 30-JAN-2003.

XX PF 29-MAR-2002; 2002WO-US010189.

XX PR 06-APR-2001; 2001US-0282211P.

XX PR 07-NOV-2001; 2001US-0337017P.

XX PR 07-MAR-2002; 2002US-0363210P.

XX PA (CTL1-) CTL IMMUNOTHERAPIES CORP.

XX PI Simard JUL, Diamond DC, Liu L, Xie Z;

XX DR WPI; 2003-248010/24.

XX PT Epitope having high affinity for major histocompatibility complex class I
XX PT useful for treating an animal, evaluating immunogenicity of a vaccine or
XX PT therapeutic composition and for diagnosing a disease.

XX PS Claim 1; SEQ ID NO 599; 239pp; English.

XX CC The invention relates to an isolated epitope polypeptide that has high
XX CC affinity for major histocompatibility complex (MHC) class I, and an

CC epitope cluster comprising the polypeptide. Also disclosed is a vaccine
CC or immunotherapeutic composition containing an epitope of the invention.
CC Compositions of the invention may be used in the treatment of cancer. The
CC method can be combined with a radiation therapy, chemotherapy,
CC biochemotherapy or surgery. The composition is also useful for evaluating
CC immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC
CC-peptide complexes of the invention are useful for determining specific T
CC cell frequency. This method is useful for evaluating immunological
CC response, by performing the method prior to and subsequent to an
CC immunisation step. Compositions of the invention are useful for
CC diagnosing a disease. The current sequence represents an epitope of the
CC invention with high affinity for MHC class I.

SQ Sequence 1250 BP; 373 A; 302 C; 278 G; 297 T; 0 U; 0 Other;

Query Match 100.0%; Score 576; DB 9; Length 1250;

Best Local Similarity 100.0%; Pred. No. 1.4e-171;

Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAACGGAGACGACGCTTTTCAAGGAGAGACCCAGGAGATGCTCAAAATATCAGCAAG 60
Db 59 ATGAACGGAGACGACGCTTTTCAAGGAGAGACCCAGGAGATGCTCAAAATATCAGCAAG 118
QY 61 TTACGAAAGGCTTCGATGATATTGCCAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 119 TTACGAAAGGCTTCGATGATATTGCCAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 178
QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTCAATGACTAAA 180
Db 179 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTCAATGACTAAA 238
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTCATCGTAGTAAACGGGCTGAGACTTCCAC 240
Db 239 CTAGGTTTCAAGGTCAACCTCCACCTTTCATCGTAGTAAACGGGCTGAGACTTCCAC 298
QY 241 GGAATGATTTTGGTAAACGATCGAAACCCAGGAATCAGGTTGAACGCTCTCAGATGACT 300
Db 299 GGAATGATTTTGGTAAACGATCGAAACCCAGGAATCAGGTTGAACGCTCTCAGATGACT 358
QY 301 TTCGGCAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAGCCAGCAGAGAA 360
Db 359 TTCGGCAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAGCCAGCAGAGAA 418
QY 361 AATGGTTTGAAGGAAGTCCAGAGGATCTGGCCCAAAATGATGGAAACAGCTGTGC 420
Db 419 AATGGTTTGAAGGAAGTCCAGAGGATCTGGCCCAAAATGATGGAAACAGCTGTGC 478
QY 421 CCCCCGGGAAATCCAACTACCTTGGAGAAGATTAAACAAGACATCTGGACCCCAAAAGGGG 480
Db 479 CCCCCGGGAAATCCAACTACCTTGGAGAAGATTAAACAAGACATCTGGACCCCAAAAGGGG 538
QY 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
Db 539 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 598
QY 541 AGCGACCTCGAGAGATGAGGTAATCTCCCTCG 576
Db 599 AGCGACCTCGAGAGATGAGGTAATCTCCCTCG 634

```

RESULT 4

AAV70060

ID AAV70060 standard; cDNA; 576 BP.

XX AC AAV70060;

XX DT 01-MAR-1999 (first entry)

XX DE Human SSX4 cDNA.

XX KW SSX4; SSX gene; human; tumour associated antigen; cancer; melanoma; ss.
XX OS Homo sapiens.

QY 241 GGGAAATGATTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGTCTCAGATGACT 300
Db 332 GGGAAATGATTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGTCTCAGATGACT 391
QY 301 TTCGGCAGCCTCCAGAGAAATCTCCGGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA 360
Db 392 TTCGGCAGCCTCCAGAGAAATCTCCGGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA 451
QY 361 AATGGTTTGAAGGAAGTCCAGAGGATCTCGGACCAAAATGATGGGAAACAGCTGTGC 420
Db 452 AATGATTCGGAGGAAGTCCAGAGGATCTCGGACCAAAATGATGGGAAACAGCTGTGC 511
QY 421 CCCCCGGGAAATCCAAAGTACCTTGAGAGAAATTAACAAGACATCTGGACCCAAAGGGGG 480
Db 512 CCCCCGGGAAATCCAAAGTACCTTGAGAGAAATTAACAAGACATCTGGACCCAAAGGGGG 571
QY 481 AAACATGCTCGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATATGAGAGATC 540
Db 572 GAACATGCTCGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATATGAGAGATC 631
QY 541 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTC 575
Db 632 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTC 666

RESULT 7

ID AAD54033
AC AAD54033 standard; DNA; 766 BP.

AC AAD54033;

DT 17-JUN-2003 (first entry)

DE Human colon cancer-associated polypeptide gene, SSX-2.

XX Human; colon cancer-associated polypeptide; immune response; therapy;

KW colon cancer; gene; ds.

OS Homo sapiens.

PN WO200290986-A1.

XX 14-NOV-2002.

XX 02-MAY-2002; 2002WO-US013994.

XX 04-MAY-2001; 2001US-00849602.

XX (LUDW-) LUDWIG INST CANCER RES.

PA (CORR) CORNELL RES FOUND INC.

XX Chen Y, Old LJ, Scanlan MJ, Stockert B;

PI WPI; 2003-112003/10.

XX Diagnosing colon cancer in a subject comprises identifying colon cancer-associated polypeptides as antigens that elicit immune response in colon cancer.

XX Claim 1; Page 90-91; 122pp; English.

XX The invention relates to a method for diagnosing colon cancer in a subject which comprises identifying colon cancer-associated polypeptides as antigens that elicit immune response in colon cancer. The method is useful for diagnosing, determining onset, progression, or regression of colon cancer in a subject, or for selecting a course of treatment of a subject having or suspected of having colon cancer. The colon cancer-associated polypeptides are useful as markers for diagnosing colon cancer, and for following the course of treatment of colon cancer. The present sequence is human colon cancer-associated polypeptide gene

XX Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;

Query Match 86.2%; Score 496.6; DB 7; Length 766;
Best Local Similarity 91.5%; Pred. No. 1.5e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGACCGAGACGACGCGCTTTCGAGGAGACCCAGGAGATGCTCAAAATATCAGAGAAG 60
Db 92 ATGACCGAGACGACGCGCTTTCGAGGAGACCCAGGAGATGCTCAAAATATCAGAGAAG 151
QY 61 TTACGAAAGCGCTTCGAGATGATATGCCAAATACCTTCTTAAGAAAGAGTGGGAAAAGATG 120
Db 152 ATCCAAAAGCGCTTCGAGATGATATGCCAAATACCTTCTTAAGAAAGAGTGGGAAAAGATG 211
QY 121 AAATCTCGAGAAAATCGTCTATGTATATGAAGCTAACTATATGAGGTCTATGACTAAA 180
Db 212 AAAGCCTCGGAGAAAATCTTCTATGTATATGAAGAGAAAGTATGAGGCTATGACTAAA 271
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATCGTGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 272 CTAGGTTTCAAGGTCACCTCCACCTTTCATGTGTATTAACCGGGCCGAAGACTTCCAG 331
QY 241 GGGAAATGATTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGTCTCAGATGACT 300
Db 332 GGGAAATGATTTGGTAAACGATCGAAACACAGGTTGAACGTCTCAGATGACT 391
QY 301 TTCGGCAGCCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA 360
Db 392 TTCGGCAGCCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA 451
QY 361 AATGGTTTGAAGGAAGTCCAGAGGATCTCGGACCAAAATGATGGGAAACAGCTGTGC 420
Db 452 AATGATTCGGAGGAAGTCCAGAGGATCTCGGACCAAAATGATGGGAAACAGCTGTGC 511
QY 421 CCCCCGGGAAATCCAAAGTACCTTGAGAGAAATTAACAAGACATCTGGACCCAAAGGGGG 480
Db 512 CCCCCGGGAAATCCAAAGTACCTTGAGAGAAATTAACAAGACATCTGGACCCAAAGGGGG 571
QY 481 AAACATGCTCGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATATGAGAGATC 540
Db 572 GAACATGCTCGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATATGAGAGATC 631
QY 541 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTC 575
Db 632 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTC 666

RESULT 8

ID ABQ83844
XX ABQ83844 standard; cDNA; 766 BP.

AC ABQ83844;

DT 03-FEB-2003 (first entry)

XX Human SSX-2 encoding cDNA SEQ ID NO:6.

XX Human; epitope; vaccine; immunotherapeutic; cytostatic; immunogenicity;
KW T cell; gene; ss.

OS Homo sapiens.

PN WO200281646-A2.

XX 17-OCT-2002.

XX 04-APR-2002; 2002WO-US011101.

XX 06-APR-2001; 2001US-0282211P.

PR 07-NOV-2001; 2001US-0337017P.

PR 07-WAR-2002; 2002US-0363210P.

XX (CTL1-) CTL IMMUNOTHERAPIES CORP.

XX Simard JLL, Diamond DC, Liu L, Xie Y;

272 CTAGGTTTCAAGGCCACCCCTCCACCTTTCATGTGTAATAAACGGGCGGAAGACTTCCAG 331
 241 GGGAAATGATTTGTACGATCGAATCAACAGGAAATCAGGTTGAACGCTCTCAGATGACT 300
 332 GGGAAATGATTTGGGATAATGACCTTAACCGGTGGGAATCAGGTTGAACGCTCTCAGATGACT 391
 301 TTCCGGCAGCTCCAGAGAAATCTCCCGAAGATCATGCCAGAGAACCCAGCAGAGGAAGAA 360
 392 TTCCGGCAGCTCCAGGAAATCTCCCGAAGATCATGCCAGAGAACCCAGCAGAGGAAGAA 451
 361 AATGGTTGAAGGAGTCCAGAGGATCTGGCCCAAAATATGAGGAAACAGCTGTGC 420
 452 AATGATTCGGAGGAGTCCAGAGGATCTGGCCCAAAATATGAGGAAACAGCTGTGC 511
 421 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAACAGGATCTGGACCCAAAGGGGG 480
 512 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAACAGGATCTGGACCCAAAGGGGG 571
 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAAGAGATC 540
 572 GAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAAGAGATC 631
 541 AGCAGCCTGAGGAGATGACGAGTAACCTCCCTC 575
 632 AGCAGCCTGAGGAGATGACGAGTAACCTCCCTC 666

RESULT 10

ADD25523
 ID ADD25523 standard; DNA; 766 BP.
 XX AC ADD25523;
 XX DT 15-JAN-2004 (first entry)
 XX DE Binding domain-immunoglobulin fusion protein-associated DNA #46.

ds; Binding domain; immunoglobulin; fusion protein; cytostatic;
 antiarthritic; immunosuppressive; antidiabetic; antithyroid;
 neuroprotective; hinge region; immunoglobulin heavy chain;
 CH2 constant region; CH3 constant region; IgG1; complement fixation;
 antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
 malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
 rheumatoid arthritis; myasthenia gravis; Grave's disease;
 type I diabetes mellitus; multiple sclerosis; autoimmune disease.

Unidentified.

XX OS Unidentified.

XX PN US2003118592-A1.

XX PD 26-JUN-2003.

XX PF 25-JUL-2002; 2002US-00207655.

XX PR 17-JAN-2001; 2001US-0367358P.

XX PR 17-JAN-2002; 2002US-00053530.

XX PR 03-JUN-2002; 2002US-0385691P.

XX PA (GENE-) GENE-CRAFT INC.

XX PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;

XX DR WPI; 2003-801317/75.

XX PT New binding domain-immunoglobulin fusion protein, useful for treating a
 PT subject having or suspected of having a malignant condition or a B-cell
 XX disorder, e.g. melanoma, Grave's disease or autoimmune disease.

XX PS Disclosure; SEQ ID NO 84; 157pp; English.

XX CC Unidentified

XX SQ Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;

Query Match 86.2%; Score 496.6; DB 9; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1.5e-146; Indels 0; Gaps 0;
 Matches 526; Conservative 0; Mismatches 49;
 QY 1 ATGAACGGAGACGACGCTTTTGAAGGAGAGCCAGGAGATGCTCAAAATATCAGAGAAG 60
 DB 92 ATGAACGGAGACGACGCTTTTGAAGGAGAGCCACGCTTGGTGTCTCAATACCAAGAGAAG 151
 QY 61 TTACGAAAGCCTTCGATGATATGCCAAATATCTCTTAAGAAAGATGGGAAAGATG 120
 DB 152 ATCCAAAGAGCCTTCGATGATATTTGCCAAATATCTCTTAAGAAAGATGGGAAAGATG 211
 QY 121 AAATCTCTCGAGAGAAATCGTCTATGTATATCAAGCTAACTATAGGTCTCATGACTAAA 180
 DB 212 AAAGCCTCGAGAAATCTTCTATGTATATGAGAGAAAGATATGAGGCTATGACTAAA 271
 QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATCGTAGTAAACGGGTGACAGCTTCCAC 240
 DB 272 CTAGGTTTCAAGGTCACCTCCACCTTTCATCGTAGTAAACGGGTGACAGCTTCCAC 331
 QY 241 GGGAAATGATTTGGTAAACGATCGAAACCCAGGAAATCAGGTTGAACGCTCTCAGATGACT 300
 DB 332 GGGAAATGATTTGGTAAACGATCGAAACCCAGGAAATCAGGTTGAACGCTCTCAGATGACT 391
 QY 301 TTCCGGCAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCAGAGAACCCAGCAGAGGAAGAA 360
 DB 392 TTCCGGCAGCCTCCAGGAAATCTTCCCGAAGATCATGCCAGAGAACCCAGCAGAGGAAGAA 451
 QY 361 AATGGTTTGAAGGAGTCCAGAGGATCTGGCCCAAAATATGAGGAAACAGCTGTGC 420
 DB 452 AATGATTCGGAGGAGTCCAGAGGATCTGGCCCAAAATATGAGGAAACAGCTGTGC 511
 QY 421 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAACAGGATCTGGACCCAAAGGGGG 480
 DB 512 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAACAGGATCTGGACCCAAAGGGGG 571
 QY 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAAGAGATC 540
 DB 572 GAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAAGAGATC 631
 QY 541 AGCAGCCTGAGGAGATGACGAGTAACCTCCCTC 575
 DB 632 AGCAGCCTGAGGAGATGACGAGTAACCTCCCTC 666

RESULT 11

AAT47748
 ID AAT47748 standard; cDNA; 931 BP.

XX AC AAT47748;

XX DT 05-JUN-1997 (first entry)

XX DE Melanoma cell antigen HOM-MEL 40 cDNA clone.

XX KW Melanoma; tumour antigen; HOM-MEL 40; serological fishing; vaccine; ds.

XX OS Homo sapiens.

XX FN W09640209-A1.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US009726.

XX PR 07-JUN-1995; 95US-00479328.

XX PR 03-JAN-1996; 96US-00580980.

XX PR 10-MAY-1996; 96US-00644116.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Pfreundschuh M, Rammensee H;

```

XX DR WPI; 1997-051896/05.
XX DR
XX PT Novel method of serological fishing - allows isolation of molecules (esp.
XX PT antigens) associated with pathological conditions, e.g. Hodgkin's
XX PT disease, cancer or auto-immune disease.
XX PS
XX PS Claim 31; Page 29; 44pp; English.
XX CC
XX CC A cDNA clone (AA47748) codes for a novel melanoma cell antigen
XX CC designated HOM-MEL 40. It was isolated using a novel method in which a
XX CC cDNA library prep'd. from malignant melanoma and expressed in E. coli
XX CC cells. Lysates of the host cells were screened with sera that had been
XX CC treated to remove interfering binding partners. This involved contacting
XX CC the sample with lysates of untransfected host cells and with host cells
XX CC transformed with the same vector (phage lambda) used to make the cDNA
XX CC library. The method, termed serological fishing, can be used to detect
XX CC antigens in human tissues, esp. tumour cells, which are useful in the
XX CC molecular diagnosis of diseases and/or for immunotherapy and gene therapy
XX CC of infectious, autoimmune and malignant diseases (see also AA47747-49).
XX CC HOM-MEL 40 is expressed in 50% of melanomas, 20% of prostate cancers, 20%
XX CC of gastric cancers, 26% of colorectal cancers, 12% of lung cancers and
XX CC 20% of breast cancers, but not in normal tissue. HLA-A2 positive tumour
XX CC cells present a nonmer (see also AA09449- 52) derived from HOM-MEL,
XX CC suggesting that HOM-MEL 40-specific vaccines, useful in inducing
XX CC cytotoxic T lymphocytes, are possible
XX SQ
XX SQ Sequence 931 BP; 274 A; 207 C; 231 G; 219 T; 0 U; 0 Other;

Query Match      86.2%; Score 496.6; DB 2; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.6e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGACCGGAGCAGCGCTTGCAGGAGAGCCAGGATGATGCTCAAAATATCAGAGAAG 60
DB 99 ATGACCGGAGCAGCGCTTGCAGGAGAGCCAGGATGATGCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGCGAAAGATG 120
DB 159 ATCCAAAGGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGCGAAAGATG 218
QY 121 AAATCTCGGAGAAATCTGCTATGTATATGAGCTAAACTATGAGTCTATGACTTAAA 180
DB 219 AAAGCTCGGAGAAATCTCTATGTATATGAGGAGAAAGATGAGGCTATGACTTAAA 278
QY 181 CTAGGTTTCAAGTCAACCTCCACCTTCCAGTAAAGCTAACTAGAGTCTATGACTTAAA 278
QY 241 GGGATGATTTGGTAAAGATCGAAACCAAGGATCAAGTTGAACGTCCTCAGATGACT 300
DB 339 GGGATGATTTGGTAAATGATGACCTTAACCGTGGGAATCAGGTTGAACGTCCTCAGATGACT 398
QY 301 TTGGCGACCTCCAGAGATCTTCCGAGATCATGCCCAAGAGCCAGAGAGAGAGAA 360
DB 399 TTGGCGACCTCCAGAGATCTTCCGAGATCATGCCCAAGAGCCAGAGAGAGAGAA 458
QY 361 AATGTTTGAAGGAGTCCAGAGGATCTGGCCCAAAATATGAGGAAAGAGCTGTGC 420
DB 459 AATGTTTGAAGGAGTCCAGAGGATCTGGCCCAAAATATGAGGAAAGAGCTGTGC 518
QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTTAAACAAGATCTTGGACCCCAAAAGGGGG 480
DB 519 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTTCAAGAGATCTTGGACCCCAAAAGGGGG 578
QY 481 AAACATGCTGAGACCAAGATCTGGTGGAGAGAAAGAGCTGGTGGTTTATGAAGAGATC 540
DB 579 GAACATGCTGAGACCAAGATCTGGTGGAGAGAAAGAGCTGGTGGTTTATGAAGAGATC 638
QY 541 AGCGACCTTGAGAGATGACGAGTAACCTCCCTC 575
DB 639 AGCGACCTTGAGAGATGACGAGTAACCTCCCTC 673

```

```

RESULT 12
AAV04267
ID AAV04267 standard; cDNA; 931 BP.
XX AC AAV04267;
XX DT 22-JUN-1998 (first entry)
XX DE Melanoma antigen clone HOM-MEL-40.
XX KW Melanoma; marker; antigen; diagnosis; serological fishing; human;
XX KW HOM-MEL-40; ss.
XX OS Homo sapiens.
XX PN W09748721-A1.
XX PD 24-DEC-1997.
XX PF 23-JUN-1997; 97WO-US010926.
XX PR 21-JUN-1996; 96US-00668128.
XX PA (LUDM-) LUDWIG INST CANCER RES.
XX PI Pfreundschuh M;
XX WPI; 1998-063074/06.
XX Nucleic acid and derived protein are markers for Hodgkin's disease - used
XX PT in identifying immuno-reactive markers of disease.
XX PS Example 7; Page 30-31; 47pp; English.
XX CC Clone HOM-MEL-40 was isolated from a human malignant melanoma cDNA
XX CC library using a method, designated serological fishing, designed to
XX CC identify immunoreactive markers of disease. In this method, cells
XX CC characteristic of a disease are used to prepare a cDNA library for
XX CC transformation of eukaryotic or prokaryotic cells, and the cells grown to
XX CC express proteins. Patient serum is incubated with the cells used to
XX CC prepare the library, but not transfected, to remove any components
XX CC reactive with these cells, then the stripped sample is similarly treated
XX CC with cells carrying the empty vector. The twice-stripped sample is
XX CC incubated with lysate of the library cells, so that specific components
XX CC in the sample may bind to the expression protein. Proteins that form
XX CC immune complexes are identified as disease markers. In the case of HOM-
XX CC MEL-40, the new melanoma associated antigen is strongly expressed in
XX CC melanoma, but not healthy tissues. The deduced amino acid sequence
XX CC includes 3 tumour-associated peptides (see AA41587-89) that bind to HLA-
XX CC A2.1. The serological fishing method was also used to identify a claimed
XX CC marker (see AAV04262) of Hodgkin's disease
XX SQ
XX SQ Sequence 931 BP; 274 A; 207 C; 231 G; 219 T; 0 U; 0 Other;

Query Match      86.2%; Score 496.6; DB 2; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.6e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGACCGGAGCAGCGCTTGCAGGAGAGCCAGGATGATGCTCAAAATATCAGAGAAG 60
DB 99 ATGACCGGAGCAGCGCTTGCAGGAGAGCCAGGATGATGCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGCGAAAGATG 120
DB 159 ATCCAAAGGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGCGAAAGATG 218
QY 121 AAATCTCGGAGAAATCTGCTATGTATATGAGCTAAACTATGAGTCTATGACTTAAA 180
DB 219 AAAGCTCGGAGAAATCTCTATGTATATGAGGAGAAAGATGAGGCTATGACTTAAA 278
QY 181 CTAGGTTTCAAGTCAACCTCCACCTTCCAGTAAAGCTAACTAGAGTCTATGACTTAAA 278

```


QY 241 GGAATGATTTTGGTAACGATGAAACACAGGAATCAGTTTGAACGCTCTCAGATGACT 300
 Db 339 GGAATGATTTTGGTAACGATGAAACACAGGAATCAGTTTGAACGCTCTCAGATGACT 398
 QY 301 TTGGCAGCCTCCAGAGAATCTTCCGGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA 360
 Db 399 TTGGCAGCCTCCAGGAATCTCCCGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA 458
 QY 361 AATGTTTGAAGAATGCCAGAGGATCTGGCCCAAAATATGGGAACAGCTGTGC 420
 Db 459 AATGATTCGGAGAATGCCAGAGGATCTGGCCCAAAATATGGGAACAGCTGTGC 518
 QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAAGAGATCTGGACCCAAAAGGGG 480
 Db 519 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAAGAGATCTGGACCCAAAAGGGG 578
 QY 481 AATGATGCTGGACCCAGAGATCTGGCCCAAAATATGGGAACAGCTGTGC 540
 Db 579 GAACATGCTGGACCCAGAGATCTGGCCCAAAATATGGGAACAGCTGTGC 638
 QY 541 AGCGACCTCAGGAAGATGAGGATTAACCTCCCTC 575
 Db 639 AGCGACCTCAGGAAGATGAGGATTAACCTCCCTC 673
 RESULT 14
 ABL63730
 ID ABL63730 standard; DNA; 1309 BP.
 XX AC ABL63730;
 XX DT 15-MAY-2002 (first entry)
 XX DE Breast cancer related gene sequence SEQ ID NO:2067.
 XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX OS Homo sapiens.
 XX PN WC200194629-A2.
 XX PD 13-DEC-2001.
 XX PF 30-MAY-2001; 2001WO-US010838.
 XX 05-JUN-2000; 2000US-0209473P.
 PR 05-JUN-2000; 2000US-0209531P.
 PR 18-SEP-2000; 2000US-0233133P.
 PR 18-SEP-2000; 2000US-0233617P.
 PR 20-SEP-2000; 2000US-0234009P.
 PR 20-SEP-2000; 2000US-0234034P.
 PR 20-SEP-2000; 2000US-0234052P.
 PR 22-SEP-2000; 2000US-0234509P.
 PR 22-SEP-2000; 2000US-0234567P.
 PR 25-SEP-2000; 2000US-0234923P.
 PR 25-SEP-2000; 2000US-0234924P.
 PR 25-SEP-2000; 2000US-0235077P.
 PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 27-SEP-2000; 2000US-0235863P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.

PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 PA (AVAL-) AVALON PHARM.
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 DR Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX Claim 1; SEQ ID NO 2067; 44pp; English.
 XX The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC activity is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 CC tumour
 XX Sequence 1309 BP; 385 A; 329 C; 281 G; 309 T; 0 U; 5 Other;
 SQ Query Match 86.2%; Score 496.6; DB 6; Length 1309;
 Best Local Similarity 91.5%; Pred. No. 1.9e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 1 ATGACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
 Db 99 ATGACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 158
 QY 61 TTACGAAAGGCTTCGATGATATGTCCTCAATATGTCCTTAAGAAAGAGTGGGAAAGATG 120
 Db 159 ATCCAAAGGCTTCGATGATATGTCCTCAATATGTCCTTAAGAAAGAGTGGGAAAGATG 218
 QY 121 AAATCTCTCGAGAAATCGTCTATGTTATATGAAGCTAACTATGAGTCTGACTATAA 180
 Db 219 AAAGCTCGAGAAATCTCTATGTTATGTAAGAGAAAGATGATGAGGCTATGACTAAA 278
 QY 181 CTAGGTTTCAAGTCCACCTCCACCTTTCATCGGTAGTAAACGGGTGCGAGCTTCCAC 240
 Db 279 CTAGGTTTCAAGGCTCCACCTCCACCTTTCATGTAATAAACCGGCCGAGACTTCCAG 338
 QY 241 GGAATGATTTTGGTAACGATGAAACACAGGAATCAGTTTGAACGCTCTCAGATGACT 300

Db 339 GGGATGATTGGTAATGACCTAACCGTGGGAATCAGTTGAACGTCCTCAGATGACT 398
Qy 301 TTGGGAGCTCCAGAGATCTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAAGA 360
Db 399 TTGGGAGCTCCAGGGAATCTCCCGAGATCATGCCCAAGAGCCAGCAGAGGAAGA 458
Qy 361 AATGGTTTGAAGGAAGTCCAGAGGATCTGGGCCACAAATGATGGGAAACAGCTGTGC 420
Db 459 AATGATTCGGAGAGTGCAGAGATCTGGGCCCAAAATGATGGGAAAGAGCTGTGC 518
Qy 421 CCCCCGGGAAATCCAGTACTCTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
Db 519 CCCCCGGGAAACCAACTCTCTGAGAGATTCACGAGAGATCTGGACCCAAAGGGGG 578
Qy 481 AATATCCTCTGGACCCACAGCTCGCTGAGAGAGAGAGAGAGAGAGAGATC 540
Db 579 GAACATCCTCTGGACCCACAGCTCGCTGAGAGAGAGAGAGAGAGAGATC 638
Qy 541 AGCCACCTCGAGGAGATGACGAGTAACTCCCTC 575
Db 639 AGCCACCTCGAGGAGATGACGAGTAACTCCCTC 673

RESULT 15
AAS80402
ID AAS80402 standard; cDNA; 921 BP.
XX
AC AAS80402;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding: novel human diagnostic protein #16206.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PSDB; ABG16215.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 16206; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 921 BP; 267 A; 205 C; 233 G; 216 T; 0 U; 0 Other;

Query Match 84.6%; Score 487.2; DB 5; Length 921;
Best Local Similarity 91.3%; Pred. No. 1.6e-143;
Matches 528; Conservative 0; Mismatches 48; Indels 2; Gaps 1;
Qy 1 ATCAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAATATATCAGAGAAG 60
Db 54 ATCAACGGAGATGACACCTTTGCAAGGAGACCCAGGAGATGATCTCAATATATCAGAGAAG 113
Qy 61 TTACGAAAGGCTTTCGATGATATTGCCAAATATCTTCTTAAGAAAGAGTGGGAAGATG 120
Db 114 ATCAAAAGGCTTTCGATGATATTGCCAAATATCTTCTTAAGAAAGAGTGGGAAGATG 173
Qy 121 AATCTCTCGAGAGAAATCGTCTATGTATATGAAGTAACTATGAGGTCTGACTTAA 180
Db 174 AAGTCTCGGAGAAATCGTCTATGTATATGAAGTAACTATGAGGTCTGACTTAA 233
Qy 181 CTAGGTTTCAAGTTCACCTTCCACCTTTCATGCTAGTAAACGGGCTGAGACTTCCAC 240
Db 234 CTAGGTTTCAAGGCTTCCCTCCCTTTCATGCTATATTAACGGGTCACAGACTTCCAG 293
Qy 241 GGGAAATGATTTGGTAAAGTTCGAAACCCAGAGATCAGGTTGAACTCTCTCAGATGACT 300
Db 294 GGGAAATGATTTGGTAAAGTTCGAAACCCAGGTTTCAACGCTCTCTCAGATGACT 353
Qy 301 TTGGGAGCTTCCAGAGATCTTCCCGAAGATATGCCCAGAGATGAGCCAGAGAGAGAA 360
Db 354 TTGGGAGCTTCCAGAGATCTTCCCGAAGATATGCCCAGAGATGAGCCAGAGAGAGAA 413
Qy 361 AATGTTTGAAGGAAGTCCAGAGGATCTTGGCCCAAAATGATGGGAAACAGCTGTGC 420
Db 414 AATGTTTGAAGGAAGTCCAGAGGATCTTGGCCCAAAATGATGGGAAACAGCTGTGC 473
Qy 421 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 480
Db 474 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 533
Qy 481 AATCATGCTCGACCCACAGACTCGGTGAGAGAAAGCAGCTGG--TGGTTTATGAAGAGA 538
Db 534 GAACATGCTCGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGGTGGTTTATGAAGAGA 593
Qy 539 TCAGCGACCTTGAGGAAGATCAGAGTAACTCCCTCTG 576
Db 594 TCAGCGACCTTGAGGAAGATCAGAGTAACTCCCTCTG 631

Search completed: March 31, 2004, 15:40:56
Job time : 441 secs


```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-796-780-1

Query Match      100.0%; Score 576; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGAGCGCTTTGCAAGGAGAGCCAGGGATGATGCTCAAAATATCAGAGAG 60
Db 1 ATGAACGGAGACGAGCGCTTTGCAAGGAGAGCCAGGGATGATGCTCAAAATATCAGAGAG 60
QY 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGGGAAAGATG 120
Db 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCTCGGAGAAATCGTCTATGTGATATGAGCTTAACTATGAGGTGATGACTTAA 180
Db 121 AAATCTCGGAGAAATCGTCTATGTGATATGAGCTTAACTATGAGGTGATGACTTAA 180
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGGTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGGAAATGATTTTGGTAACGATCGAATCCAGGATCAGGTTGAAACGTCCTCAGATGACT 300
Db 241 GGGAAATGATTTTGGTAACGATCGAATCCAGGATCAGGTTGAAACGTCCTCAGATGACT 300
QY 301 TTCCGAGCCTCCAGAGAAATCTCCGAGAAATCATGCCAAGAGCCAGCAGAGAGAGAA 360
Db 301 TTCCGAGCCTCCAGAGAAATCTCCGAGAAATCATGCCAAGAGCCAGCAGAGAGAGAA 360
QY 361 AATGTTTGAAGGAGTCCAGAGAGTCTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 420
Db 361 AATGTTTGAAGGAGTCCAGAGAGTCTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 420
QY 421 CCCCCGGGAAATCCAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
Db 421 CCCCCGGGAAATCCAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
QY 481 AAACATCCTGGAGCCAGAGTCCGAGAGAGTCCGAGAGAGAGAGAGATC 540
Db 481 AAACATCCTGGAGCCAGAGTCCGAGAGAGTCCGAGAGAGAGAGAGATC 540
QY 541 AGCGACCTCGAGGAGATGACGAGTAACTCCCTCTCG 576
Db 541 AGCGACCTCGAGGAGATGACGAGTAACTCCCTCTCG 576

RESULT 4
US-09-344-040C-5
; Sequence 5, Application US/09344040C
; Patent No. 6548064
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; APPLICANT: Ramensee, Hans Georg
; APPLICANT: Stavanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determin
; TITLE OF INVENTION: Expression of an SSX Gene, Peptides Derived From Said SSX Gene a
; FILE REFERENCE: LUD 5556.1
; CURRENT APPLICATION NUMBER: US/09/344,040C
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 132
; SEQ ID NO 5
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-344-040C-5

Query Match      100.0%; Score 576; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGAGCGCTTTGCAAGGAGAGCCAGGGATGATGCTCAAAATATCAGAGAG 60
Db 1 ATGAACGGAGACGAGCGCTTTGCAAGGAGAGCCAGGGATGATGCTCAAAATATCAGAGAG 60
QY 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGGGAAAGATG 120
Db 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCTCGGAGAAATCGTCTATGTGATATGAGCTTAACTATGAGGTGATGACTTAA 180
Db 121 AAATCTCGGAGAAATCGTCTATGTGATATGAGCTTAACTATGAGGTGATGACTTAA 180
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGGTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGGAAATGATTTTGGTAACGATCGAATCCAGGATCAGGTTGAAACGTCCTCAGATGACT 300
Db 241 GGGAAATGATTTTGGTAACGATCGAATCCAGGATCAGGTTGAAACGTCCTCAGATGACT 300
QY 301 TTCCGAGCCTCCAGAGAAATCTCCGAGAAATCATGCCAAGAGCCAGCAGAGAGAGAA 360
Db 301 TTCCGAGCCTCCAGAGAAATCTCCGAGAAATCATGCCAAGAGCCAGCAGAGAGAGAA 360
QY 361 AATGTTTGAAGGAGTCCAGAGAGTCTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 420
Db 361 AATGTTTGAAGGAGTCCAGAGAGTCTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 420
QY 421 CCCCCGGGAAATCCAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
Db 421 CCCCCGGGAAATCCAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
QY 481 AAACATCCTGGAGCCAGAGTCCGAGAGAGTCCGAGAGAGAGAGATC 540
Db 481 AAACATCCTGGAGCCAGAGTCCGAGAGAGTCCGAGAGAGAGAGATC 540
QY 541 AGCGACCTCGAGGAGATGACGAGTAACTCCCTCTCG 576
Db 541 AGCGACCTCGAGGAGATGACGAGTAACTCCCTCTCG 576

RESULT 5
US-09-833-039A-5
; Sequence 5, Application US/09833039A
; Patent No. 6673350
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: US/09/833,039A
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 09/409,455
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/344,040
; PRIOR FILING DATE: 1999-06-26
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-09-833-039A-5

Query Match 100.0%; Score 576; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 ATGAACGGAGACGACGCTTTGCAAGGAGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
Db 1 ATGAACGGAGACGACGCTTTGCAAGGAGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60

Qy 61 TTACGAAGGCTTCGATGATATTGCCAAATACCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 61 TTACGAAGGCTTCGATGATATTGCCAAATACCTTCTTAAGAAAGAGTGGGAAAGATG 120

Qy 121 AAATCCTCGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTAA 180
Db 121 AAATCCTCGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTAA 180

Qy 181 CTAGGTTTCAAGGTCACTCCACCTTTCATGCGTAGTAAACCGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGGTCACTCCACCTTTCATGCGTAGTAAACCGGCTGCAGACTTCCAC 240

Qy 241 GGGATGATTTTGGTAAAGTCGAAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Db 241 GGGATGATTTTGGTAAAGTCGAAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300

Qy 301 TTCCGACGCTCCAGAGAAATCTTCCGAAAGATCATGCCAAGAGCCAGCAGAGAGAGAA 360
Db 301 TTCCGACGCTCCAGAGAAATCTTCCGAAAGATCATGCCAAGAGCCAGCAGAGAGAGAA 360

Qy 361 AATGGTTGAAGGAGTCCAGAGCATCTCGGCCAACAATAATGATGGGAAACAGCTGTGC 420
Db 361 AATGGTTGAAGGAGTCCAGAGCATCTCGGCCAACAATAATGATGGGAAACAGCTGTGC 420

Qy 421 CCCCCGGGAAATCAAAGTACCTTGGAGAGATTAACAGAGATCTGGACCCCAAGAGGGG 480
Db 421 CCCCCGGGAAATCAAAGTACCTTGGAGAGATTAACAGAGATCTGGACCCCAAGAGGGG 480

Qy 481 AAACATGCTGGACCCACAGACTCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
Db 481 AAACATGCTGGACCCACAGACTCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540

Qy 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576
Db 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576

```

RESULT 6
US-09-392-714-17
; Sequence 17, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: L0461/7062
; CURRENT FILING DATE: 1999-09-09
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-392-714-17

Query Match 100.0%; Score 576; DB 4; Length 576;

Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 ATGAACGGAGACGACGCTTTGCAAGGAGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
Db 1 ATGAACGGAGACGACGCTTTGCAAGGAGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60

Qy 61 TTACGAAGGCTTCGATGATATTGCCAAATACCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 61 TTACGAAGGCTTCGATGATATTGCCAAATACCTTCTTAAGAAAGAGTGGGAAAGATG 120

Qy 121 AAATCCTCGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTAA 180
Db 121 AAATCCTCGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTAA 180

Qy 181 CTAGGTTTCAAGGTCACTCCACCTTTCATGCGTAGTAAACCGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGGTCACTCCACCTTTCATGCGTAGTAAACCGGCTGCAGACTTCCAC 240

Qy 241 GGGATGATTTTGGTAAAGTCGAAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Db 241 GGGATGATTTTGGTAAAGTCGAAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300

Qy 301 TTCCGACGCTCCAGAGAAATCTTCCGAAAGATCATGCCAAGAGCCAGCAGAGAGAGAA 360
Db 301 TTCCGACGCTCCAGAGAAATCTTCCGAAAGATCATGCCAAGAGCCAGCAGAGAGAGAA 360

Qy 361 AATGGTTGAAGGAGTCCAGAGCATCTCGGCCAACAATAATGATGGGAAACAGCTGTGC 420
Db 361 AATGGTTGAAGGAGTCCAGAGCATCTCGGCCAACAATAATGATGGGAAACAGCTGTGC 420

Qy 421 CCCCCGGGAAATCAAAGTACCTTGGAGAGATTAACAGAGATCTGGACCCCAAGAGGGG 480
Db 421 CCCCCGGGAAATCAAAGTACCTTGGAGAGATTAACAGAGATCTGGACCCCAAGAGGGG 480

Qy 481 AAACATGCTGGACCCACAGACTCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
Db 481 AAACATGCTGGACCCACAGACTCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540

Qy 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576
Db 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576

```

RESULT 7
US-09-392-714-16
; Sequence 16, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: L0461/7062
; CURRENT FILING DATE: 1999-09-09
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 766
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-392-714-16

Query Match 86.2%; Score 496.6; DB 4; Length 766;
Best Local Similarity 91.5%; Pred. No. 1.3e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCGCTTTCAGAGGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 92 ATGAACGAGACGACGCGCTTTCAGAGGACCCAGGATGATGCTCAATATCAGAGAAG 151
QY 61 TTACGAAAGGCGCTTCGATGATATGCAAAATCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 152 ATCCAAAGGCGCTTCGATGATATGCAAAATCTTCTTAAGAAAGAGTGGGAAAGATG 211
QY 121 AAATCTCGGAGAAATCTCTATGATATGCAAGCTTAACTATGAGTCTGACTTAA 180
Db 212 AAAGCCTCGGAGAAATCTCTATGATATGCAAGCTTAACTATGAGTCTGACTTAA 271
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTCATGCGTAGTAAACGGGCTCGAGACTTCCAC 240
Db 272 CTAGGTTTCAAGGCAACCTCCACCTTTCATGCTGTAATTAACGGGCGGAGACTTCCAG 331
QY 241 GGAATGATTTTGGTAAACGATGAAACCCACAGAAATCAGGTTGAACTCTCTGATGACT 300
Db 332 GGAATGATTTTGGTAAACGATGAAACCCACAGAAATCAGGTTGAACTCTCTGATGACT 391
QY 301 TTCCGAGGCTCCAGAGATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGGAAGAA 360
Db 392 TTCCGAGGCTCCAGAGATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGGAAGA 451
QY 361 AATGTTTGAAGGAGTGGCCAGAGGATCTTGGCCCAAAATGATGGGAAACAGCTGTGC 420
Db 452 AATGTTTGGAGGAGTGGCCAGAGGATCTTGGCCCAAAATGATGGGAAAGAGCTGTGC 511
QY 421 CCCCCGGGAAATCCAAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 480
Db 512 CCCCCGGGAAATCCAAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 571
QY 481 AAACATGCTGGACCCACAGACTGCTGAGAGAAAGCAGCTGCTGTTTATGAGAGATC 540
Db 572 GAACATGCTGGACCCACAGACTGCTGAGAGAAAGCAGCTGCTGTTTATGAGAGATC 631
QY 541 AGCGACCTGAGGAAGATGACGAGTAACTCCCTC 575
Db 632 AGCGACCTGAGGAAGATGACGAGTAACTCCCTC 666

RESULT 8

US-08-479-328-2
; Sequence 2, Application US/08479328
; Patent No. 5698396
; GENERAL INFORMATION:
; APPLICANT: Pfeundschuh, Michael
; TITLE OF INVENTION: Method For Identifying Or Isolating A Molecule
; TITLE OF INVENTION: And Molecules Identified Thereby
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,328
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5698396man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5410
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 931 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
US-08-479-328-2

Query Match 86.2%; Score 496.6; DB 1; Length 931;

Best Local Similarity 91.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCGCTTTCAGAGGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 99 ATGAACGAGACGACGCGCTTTCAGAGGACCCAGGATGATGCTCAATATCAGAGAAG 158
QY 61 TTACGAAAGGCGCTTCGATGATATGCAAAATCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 159 ATCCAAAGGCGCTTCGATGATATGCAAAATCTTCTTAAGAAAGAGTGGGAAAGATG 218
QY 121 AAATCTCGGAGAAATCTCTATGATATGCAAGCTTAACTATGAGTCTGACTTAA 180
Db 219 AAAGCCTCGGAGAAATCTCTATGATATGCAAGCTTAACTATGAGTCTGACTTAA 278
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTCATGCGTAGTAAACGGGCTCGAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGCAACCTCCACCTTTCATGCTGTAATTAACGGGCGGAGACTTCCAG 338
QY 241 GGAATGATTTTGGTAAACGATGAAACCCACAGAAATCAGGTTGAACTCTCTGATGACT 300
Db 339 GGAATGATTTGGATTAATGACCTTAACCTGGGAAATCAGTTGAACTCTCTGATGACT 398
QY 301 TTCCGAGGCTCCAGAGATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGGAAGAA 360
Db 399 TTCCGAGGCTCCAGGAAATCTCCCGAAGATCATCCCAAGAGCCAGCAGAGGAAGA 458
QY 361 AATGTTTGAAGGAGTGGCCAGAGGATCTGCCCCAATAATGATGGGAAACAGCTGTGC 420
Db 459 AATGATTCGGAGGAGTGGCCAGAGGATCTGCCCCAATAATGATGGGAAAGAGCTGTGC 518
QY 421 CCCCCGGGAAATCCAAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 480
Db 519 CCCCCGGGAAACCAACTACCTCTGAGAGATTCACGAGAGATCTGGACCCCAAGGGGG 578
QY 481 AAACATGCTGGACCCACAGACTGCTGAGAGAAAGCAGCTGCTGTTTATGAGAGATC 540
Db 579 GAACATGCTGGACCCACAGACTGCTGAGAGAAAGCAGCTGCTGTTTATGAGAGATC 638
QY 541 AGCGACCTGAGGAAGATGACGAGTAACTCCCTC 575
Db 639 AGCGACCTGAGGAAGATGACGAGTAACTCCCTC 673

RESULT 9

US-08-761-119-2
; Sequence 2, Application US/08761119
; Patent No. 5798264
; GENERAL INFORMATION:

; APPLICANT: Pfeundschuh, Michael
; TITLE OF INVENTION: Method For Identifying Or Isolating A Molecule
; TITLE OF INVENTION: And Molecules Identified Thereby
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/761,119
FILING DATE: 6-DECEMBER-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5798264man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5410.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 931 base pairs
TYPE: nucleic acid
STRANDEDNESS: double-stranded
TOPOLOGY: linear
US-08-761-119-2

Query Match 86.2%; Score 496.6; DB 1; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
DB 99 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158

QY 61 TTACGAAGGCTTCGATGATATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 120
DB 159 ATCAAAAGGCTTCGATGATATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 218

QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCATGACTTAA 180
DB 219 AAAGCTTCGAGAAATCTTCTATGTATATGAAGAAAGTATGAGGTCATGACTTAA 278

QY 181 CTAGGTTTCAAGTTCACCTCCACCTTTCAAGGATGATGCTCAAAATATGCTCAAAATATGCTCAAA 240
DB 279 CTAGGTTTCAAGGTCACCTCCACCTTTCAAGGATGATGCTCAAAATATGCTCAAAATATGCTCAAA 338

QY 301 TTCCGACGCTCCAGAGATCTTCCGAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAA 360
DB 399 TTCCGAGGCTCCAGAGATCTTCCGAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAA 458

QY 361 AATGTTTGAAGGATGCTCCAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 420
DB 459 AATGTTTGAAGGATGCTCCAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 518

QY 421 CCCCCGGGAAATCAAGTACCTTCGAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 480
DB 519 CCCCCGGGAAATCAAGTACCTTCGAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 578

QY 481 AAACATGCTCGAGACCCAGAGATGCTCCGAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 540
DB 579 GAACATGCTCGAGACCCAGAGATGCTCCGAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 638

QY 541 AGCGACCTTCGAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 575
DB 639 AGCGACCTTCGAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 673

RESULT 10
US-08-668-128B-2
Sequence 2, Application US/08668128B
Patent No. 5840568
GENERAL INFORMATION:
APPLICANT: Pfeunderschuh, Michael
TITLE OF INVENTION: Hodgkin's Disease Associated Molecules And

TITLE OF INVENTION: Uses Thereof
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/668,128B
FILING DATE: 21-JUNE-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/644,116
FILING DATE: 10-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/580,980
FILING DATE: 03-JANUARY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,328
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5840568man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5441
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 931 base pairs
TYPE: nucleic acid
STRANDEDNESS: double-stranded
TOPOLOGY: linear
US-08-668-128B-2

Query Match 86.2%; Score 496.6; DB 2; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
DB 99 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158

QY 61 TTACGAAGGCTTCGATGATATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 120
DB 159 ATCAAAAGGCTTCGATGATATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 218

QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCATGACTTAA 180
DB 219 AAAGCTTCGAGAAATCTTCTATGTATATGAAGAAAGTATGAGGTCATGACTTAA 278

QY 181 CTAGGTTTCAAGTTCACCTCCACCTTTCAAGGATGATGCTCAAAATATGCTCAAAATATGCTCAAA 240
DB 279 CTAGGTTTCAAGGTCACCTCCACCTTTCAAGGATGATGCTCAAAATATGCTCAAAATATGCTCAAA 338

QY 241 GGGATGATTTTGGTACGATCGAAACACAGGATGATGCTCAAAATATGCTCAAAATATGCTCAAA 300
DB 339 GGGATGATTTTGGTACGATCGAAACACAGGATGATGCTCAAAATATGCTCAAAATATGCTCAAA 398

QY 301 TTCCGACGCTCCAGAGATCTTCCGAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 360
DB 399 TTCCGAGGCTCCAGAGATCTTCCGAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 458

QY 361 AATGTTTGAAGGATGCTCCAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 420
DB 459 AATGTTTGAAGGATGCTCCAGAGATGATGCTCAAAATATGCTCAAAATATGCTCAAAATATGCTCAAA 518

QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAAGATTAAACAGACATCTGGACCCCAAAAGGGG 480
| | | | |
Db 519 CCCCCGGGAAACCACTACTCTGAGAGATTCAGAGAGATCTGGACCCCAAAAGGGG 578
| | | | |
QY 481 AAATGCTTGGACCCACAGATGCTGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
| | | | |
Db 579 GAACATGCTGGACCCACAGATGCTGTGAGAGAAACAGCTGGTGTATGAGAGATC 638
| | | | |
QY 541 AGCGACCTTGAGAGATGACGAGTAACCTCCCTC 575
| | | | |
Db 639 AGCGACCTTGAGAGATGACGAGTAACCTCCCTC 673
| | | | |

RESULT 11

US-08-905-445-2
; Sequence 2, Application US/08905445
; Patent No. 5864015
; GENERAL INFORMATION:
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Hodgkin's Disease Associated Molecules And
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,445
; FILING DATE: 04-AUG-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/668,128
; FILING DATE: 21-JUNE-1996
; APPLICATION NUMBER: 08/644,116
; FILING DATE: 10-MAY-1996
; APPLICATION NUMBER: 08/580,980
; FILING DATE: 03-JANUARY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,328
; FILING DATE: 07-JUNE-1995

ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5864015man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5441
; TELEPHONE: (212) 688-9200
; TELECOMMUNICATION INFORMATION:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 931 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; US-08-905-445-2

Query Match 86.2%; Score 496.6; DB 2; Length 931;
Best Local Similarity 93.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATACAGAAAG 60
| | | | |
Db 99 ATGAACGGAGACGACGGCTTTGCAAGGAGACCCAGGTTGGTGTCTCAATACAGAAAG 158
| | | | |
QY 61 TTACAAAGGCTTCGATGATATTGCCAAATCTCTAGAAAGATGGGAAAGATG 120
| | | | |

Db 159 ATCCAAAGGCTTCGATGATATTGCCAAATCTTCTTAAGAAAGAGTGGGAAAGATG 218
| | | | |
QY 121 AAATCCTCGGAAATCGTCTATGTATATGAAGCTAAACTATGAGTCTATGACTAAA 180
| | | | |
Db 219 AAAGCTCGGAAATCTTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 278
| | | | |
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTCATGCGTAGTAAACGGGCTGAGACTTCCAC 240
| | | | |
Db 279 CTAGGTTTCAAGGCCACCTCCACCTTTCATGTGTATATAAAGCGGCCGGAAGACTTCCAG 338
| | | | |
QY 241 GGGATGATTTGGTAAACGATCGAACCACAGGATCAGGTTGAACGTCTCTCAGATGACT 300
| | | | |
Db 339 GGGATGATTTGGATTAATGACCTTAACCTGGGAATCAGGTTGAACGTCTCTCAGATGACT 398
| | | | |
QY 301 TTCCGCGAGCTCCAGAGAAATCTTCCCGAAGATCATGCCCAAGAACCCAGCAGAGGAAGA 360
| | | | |
Db 399 TTCCGCGAGCTCCAGGGAATCTCCCGAAGATCATGCCCAAGAACCCAGCAGAGGAAGA 458
| | | | |
QY 361 AATGTTTGAAGAAAGTGCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
| | | | |
Db 459 AATGATTCGGAAGTGCAGAAAGCATCTGGCCCAAAATGATGGGAAAGAGCTGTGC 518
| | | | |
QY 421 CCCCCGGGAAATCCCAAGTACCTTGGAGAAGATTAAACAGACATCTCGACCCCAAGGGGG 480
| | | | |
Db 519 CCCCCGGGAAACCACTACTCTGTGAGAAGATTACAGAGATCTCGACCCCAAGGGGG 578
| | | | |
QY 481 AAACATGCTCGACCCACAGATCTGCTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
| | | | |
Db 579 GAACATGCTCGACCCACAGATCTGCTGAGAGAAACAGCTGGTGTATGAGAGATC 638
| | | | |
QY 541 AGCGACCTTGAGAGATGACGAGTAACCTCCCTC 575
| | | | |
Db 639 AGCGACCTTGAGAGATGACGAGTAACCTCCCTC 673
| | | | |

RESULT 12

US-08-959-625-2
; Sequence 2, Application US/08959625
; Patent No. 6017716
; GENERAL INFORMATION:

APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Method For Identifying Or Isolating A Molecule
; TITLE OF INVENTION: And Molecules Identified Thereby
; NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/959,625
; FILING DATE:

CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/580,980
; FILING DATE: 03-JANUARY-1996
; APPLICATION NUMBER: 08/479,328
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6017716man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5410.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

339	GGGAATGATTGGATAATGACCTTAACCGTGGGAATCAGGTTGAACGTCTCTCAGATGACT	398
301	TTCCGCAGCTCCAGAGAAATCTTCCGAAGATCATGCCAAGAAGCCACGACAGAGAGAA	360
399	TTCCGCAGGCTCCAGGGAATCTCTCCGAGANTCATGCCAAGAAGCCACGACAGAGAGA	458
361	AATGTTTGAAGAGTCCAGAGGCATCTGGCCCAAAATGATGGGAAAACAGCTGTGC	420
459	AATGATTGGAGGAAGTCCAGAGACATCTGGGCCCAAAAATGATGGGAAAGAGCTGTGC	518
421	CCCCCGGGAATCCAAGTACCTTCGGAGAGATTAAACAAGACATCTGGACCCCAAAAGCGGG	480
519	CCCCCGGGAACCAACTACCTCTGGAAGATTACGAGAGATCTCGACCCCAAAAGCGGG	578
481	AAACATGCTCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTGTTTATGAAGAGATC	540
579	GAAATGCTTGACCCACAGACTCGGTGAGAGAAACAGCTGGTGATTATTAAGAGATC	638
541	AGCGACCTTGAGGAAGATGACGATTAATCTCCCTC	575
639	AGCGACCTTGAGGAAGATGACGATTAATCTCCCTC	673

Search completed: March 31, 2004, 17:26:48
Job time : 99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 16:39:19 ; Search time 389 Seconds
(without alignments)
5512.734 Million cell updates/sec

Title: US-09-975-856-1
Perfect score: 576
Sequence: 1 ATGACGAGACGACGCTT.....ATGACGAGTAACTCCCTCG 576

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	576	100.0	576	10	US-09-975-856-1
2	576	100.0	576	10	US-09-833-039-5
3	576	100.0	576	14	US-10-177-277-5
4	576	100.0	576	15	US-10-177-277-5
5	496.6	86.2	766	10	US-10-117-937-599
6	496.6	86.2	766	10	US-09-849-602-13
7	496.6	86.2	766	15	US-10-207-655-84
8	496.6	86.2	931	10	US-10-117-937-6
9	496.6	86.2	931	14	US-09-833-039-2
10	496.6	86.2	1309	9	US-10-177-277-2
11	496.6	86.2	1309	9	US-09-954-531-1000
12	496	86.1	1249	15	US-09-954-531-1392
13	482.2	83.7	576	10	US-10-159-563-402
14	482.2	83.7	576	10	US-09-975-856-2
15	482.2	83.7	576	14	US-09-833-039-6
					Sequence 1, Appli
					Sequence 5, Appli
					Sequence 599, App
					Sequence 13, Appl
					Sequence 84, Appl
					Sequence 6, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 1000, Ap
					Sequence 1392, Ap
					Sequence 402, App
					Sequence 2, Appli
					Sequence 6, Appli

16	480	83.3	766	10	US-09-833-039-1	Sequence 1, Appli
17	480	83.3	766	14	US-10-177-277-1	Sequence 1, Appli
C 18	308.6	53.6	574	15	US-10-027-632-106782	Sequence 106782,
C 19	308.6	53.6	718	15	US-10-027-632-24594	Sequence 24594, A
C 20	140	24.3	564	14	US-10-029-386-10234	Sequence 10234, A
C 21	137	23.8	189	14	US-10-029-386-23851	Sequence 23851, A
C 22	111.4	19.3	675	15	US-10-027-632-33096	Sequence 33096, A
C 23	111.4	19.3	675	15	US-10-027-632-33097	Sequence 33097, A
C 24	111.4	19.3	675	15	US-10-027-632-33097	Sequence 148783,
C 25	87.8	15.2	3186778	15	US-10-027-632-174961	Sequence 174961,
C 26	72	12.5	3225	15	US-10-027-632-114514	Sequence 114514,
C 27	72	12.5	3225	15	US-10-027-632-114515	Sequence 114515,
C 28	47	8.2	575	15	US-10-027-632-32431	Sequence 32431,
C 29	47	8.2	575	15	US-10-027-632-32432	Sequence 32432,
C 30	41	7.1	583	10	US-09-814-353-18480	Sequence 18480, A
C 31	40.8	7.1	60	10	US-09-908-975-6728	Sequence 6728, Ap
C 32	39.8	6.9	3591	9	US-09-981-353-84	Sequence 84, Appl
C 33	39.6	6.9	439	12	US-10-085-783A-58497	Sequence 58497, A
C 34	39.6	6.9	439	15	US-10-242-535A-58497	Sequence 58497, A
C 35	39.4	6.8	1097	14	US-10-133-013-261	Sequence 261, App
C 36	39	6.8	629	15	US-10-027-632-134770	Sequence 134770,
C 37	37.8	6.6	389	12	US-10-085-783A-18965	Sequence 18965, A
C 38	37.8	6.6	389	15	US-10-242-535A-18965	Sequence 18965, A
C 39	37.8	6.6	538	14	US-10-029-386-7375	Sequence 7375, Ap
C 40	37.8	6.6	1530	10	US-09-814-353-21781	Sequence 21781, A
C 41	37.8	6.6	2024	15	US-10-108-260A-1253	Sequence 1253, Ap
C 42	37.8	6.6	2736	13	US-10-098-841-181	Sequence 181, App
C 43	37.8	6.6	3205	14	US-10-269-909-87	Sequence 87, Appl
C 44	37.8	6.6	5532	10	US-09-971-392-98	Sequence 98, Appl
C 45	37.6	6.5	856	15	US-10-027-632-163719	Sequence 163719,

ALIGNMENTS

RESULT 1

US-09-975-856-1
; Sequence 1, Application US/09975856
; Publication No. US20030023057A1
; GENERAL INFORMATION:

APPLICANT: Gure, Ali O.; Tureci, Ozlem; Sahin, Ugur; Tsang, Solan;
Scanlan, Matthew J.; Knuth Alexander; Pfreundschuh, Michael; Old,
Lloyd J.; Chen, Yao-Tseng
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding SSX
Family
Members And Uses Thereof

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSER: Fulbright & Jaworski L.L.P.

STREET: 666 Fifth Avenue

CITY: New York City

STATE: New York

ZIP: 10103

COMPUTER TYPE: Diskette, 3.5 inch, 144 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/975,856

FILING DATE: 11-Oct-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/796,780

FILING DATE: 2001-03-01

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. US20030023057Alman D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5480

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 318-3168

TELEFAX: (212) 752-5958

INFORMATION FOR SEQ ID NO: 1:

```

/
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 576 nucleotides
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-975-856-1

Query Match          100.0%; Score 576; DB 10; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.8e-174;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1  ATGAACGGAGACGACGCGCTTGGCAAGGAGACCCAGGAGTGTCTCAATATACAGAAAG 60
Db      1  ATGAACGGAGACGACGCGCTTGGCAAGGAGACCCAGGAGTGTCTCAATATACAGAAAG 60

QY      61  TTACGAAAGGGCTTCGATGATATTGCCAAATACITTTCTTAGAAGAAGTGGGAAAGATG 120
Db      61  TTACGAAAGGGCTTCGATGATATTGCCAAATACITTTCTTAGAAGAAGTGGGAAAGATG 120

QY      121  AAATCCTCGGAGAAATCGTCTATGTGTATATGAAGCTAAACTATGAGTGCATGACTAAA 180
Db      121  AAATCCTCGGAGAAATCGTCTATGTGTATATGAAGCTAAACTATGAGTGCATGACTAAA 180

QY      181  CTAGGTTTCAAGGTCACCCCTCCACCTTTTCANGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db      181  CTAGGTTTCAAGGTCACCCCTCCACCTTTTCANGCGTAGTAAACGGGCTGCAGACTTCCAC 240

QY      241  GGGAAATGATTTTGGTAAAGTCGAAACCAACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Db      241  GGGAAATGATTTTGGTAAAGTCGAAACCAACAGGAATCAGGTTGAACGTCCTCAGATGACT 300

QY      301  TTGGCAGCCTCCAGAGAATCTTCCCGAAGATCATGCCAAGAGCCAGACAGAGGAGAA 360
Db      301  TTGGCAGCCTCCAGAGAACTCTTCCCGAAGATCATGCCAAGAGCCAGACAGAGGAGAA 360

QY      361  AATGGTTTGAAGGAAGTCCAGAGGONCTCGGCCCACAAAATGATGGAAACAGCTGTGC 420
Db      361  AATGGTTTGAAGGAAGTCCAGAGGCACTCGGCCCACAAAATGATGGAAACAGCTGTGC 420

QY      421  CCCCAGGAAATCCCACTACTCTTGGAGAAGNTAACAGAATCTGAGCCCAAAAGGGGG 480
Db      421  CCCCAGGAAATCCCACTACTCTTGGAGAAGNTAACAGAATCTGAGCCCAAAAGGGGG 480

QY      481  AATCATGCCTGGACCCACAGACTCGCTGAGAGAAAGCAGCTGCTGTTTATGAAGATC 540
Db      481  AATCATGCCTGGACCCACAGACTCGCTGAGAGAAAGCAGCTGCTGTTTATGAAGATC 540

QY      541  AGCGACCCCTGAGGAAGATGACGAGTAACCTCCCTCG 576
Db      541  AGCGACCCCTGAGGAAGATGACGAGTAACCTCCCTCG 576

```

RESULT 2
US-09-833-039-5
; Sequence 5, Application US/09833039
; Publication No. US20030175960A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Freundschuh, Michael
; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: US/09/833,039
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 09/409,455
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US 09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05

```

; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 5
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-039-5

Query Match
Best Local Similarity 100.0%; Score 576; DB 10; Length 576;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 ATGAACGAGACGACGCGCTTTGCCAAGAGACCCAGGAGTATGCTCAAATATCGAGAAG 60
Db 1 ATGAACGAGACGACGCGCTTTGCCAAGAGACCCAGGAGTATGCTCAAATATCGAGAAG 60
QY 61 TTACGAAGAGCGCTTCGATGATATTGCCAAATACTTCTTAAGAAGAGTGGGAAAAGATG 120
Db 61 TTACGAAGAGCGCTTCGATGATATTGCCAAATACTTCTTAAGAAGAGTGGGAAAAGATG 120
QY 121 AAATCCCTCGGAAATTCGTCATATGTGTATATGAAGCTAAACTATGAGGTCATCACTAAA 180
Db 121 AAATCCCTCGGAAATTCGTCATATGTGTATATGAAGCTAAACTATGAGGTCATCACTAAA 180
QY 181 CTAGGTTTCAAGGTCACCCCTCCCACTTTTCATGCGTAGTAAACGGGCTCGAGACTCCAC 240
Db 181 CTAGGTTTCAAGGTCACCCCTCCCACTTTTCATGCGTAGTAAACGGGCTCGAGACTCCAC 240
QY 241 GGGAAATGATTTTGGTAAACGATCGAAACCAAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Db 241 GGGAAATGATTTTGGTAAACGATCGAAACCAAGGAATCAGGTTGAACGTCCTCAGATGACT 300
QY 301 TTCGGCAGCCTCCAGAGAAATCTTCCGAAGATCATGCCAAGAGCCAGCAGAGGAGAA 360
Db 301 TTCGGCAGCCTCCAGAGAAATCTTCCGAAGATCATGCCAAGAGCCAGCAGAGGAGAA 360
QY 361 AATGTTTGAAGGAAAGTCCAGAGGATCTGGCCCAAAAATGATGGAAACAGCTGTGC 420
Db 361 AATGTTTGAAGGAAAGTCCAGAGGATCTGGCCCAAAAATGATGGAAACAGCTGTGC 420
QY 421 CCCCCGGGAATCCAAGTACCTTGGAGAAGATTAAACAGCATCTCGAACCCAAAGGGGG 480
Db 421 CCCCCGGGAATCCAAGTACCTTGGAGAAGATTAAACAGCATCTCGAACCCAAAGGGGG 480
QY 481 AAACATGCGCTGGACCCACAGACTCGCTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 481 AAACATGCGCTGGACCCACAGACTCGCTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
QY 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 576
Db 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 576

```

RESULT 3
US-10-177-277-5
/ Sequence 5, Application US/10177277
/ Publication No. US20030185844A1
/ GENERAL INFORMATION:
/ APPLICANT: Tureci, Ozlem
/ APPLICANT: Sahin, Ugur
/ APPLICANT: Pfreundschnuh, Michael
/ APPLICANT: Ramensee, Hans Georg
/ APPLICANT: Stevanovic, Stefan
/ TITLE OF INVENTION: Method for Determining Presence of Cancer In a Sample By Determining Expression of an SSX Gene, Peptides Derived From Said SSX Gene
/ TITLE OF INVENTION: Expression of an SSX Gene, and Uses Thereof
/ FILE REFERENCE: LUD 5556.1
/ CURRENT APPLICATION NUMBER: US/10/177,277
/ CURRENT FILING DATE: 2002-06-21
/ PRIOR APPLICATION NUMBER: US/09/344,040
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: US 09/105,839
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: US 08/851,130

```

; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 132
; SEQ ID NO 5
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-277-5

```

Query Match	100.0%	Score 576;	DB 14;	Length 576;
Best Local Similarity	100.0%;	Pred. No. 2.8e-174;		
Matches 576;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAACGAGACGACGCGCTTTGCAAGAGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG	60	
Db	1	ATGAACGAGACGACGCGCTTTGCAAGAGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG	60	
QY	61	TTACGAAGAAGGCGCTTCGATGATATTGCCAAATACATCTCTCTAAGAAGAAGTGGGAAGAAGTG	120	
Db	61	TTACGAAGAAGGCGCTTCGATGATATTGCCAAATACATCTCTCTAAGAAGAAGTGGGAAGAAGTG	120	
QY	121	AAATCTCTGGAGAAATATCGCTATATGTGTATATGAAGCTAAACTATAGAGGTCAATGACTAAA	180	
Db	121	AAATCTCTGGAGAAATATCGCTATATGTGTATATGAAGCTAAACTATAGAGGTCAATGACTAAA	180	
QY	181	CTAGGTTTCAAGGTCACCGCTCCACACCTTTATGCGTATAGTAAACGGGCTCGAGACTTCCAC	240	
Db	181	CTAGGTTTCAAGGTCACCGCTCCACACCTTTATGCGTATAGTAAACGGGCTCGAGACTTCCAC	240	
QY	241	GGGAATGATTTTGGTAAACGATCGAATCCACAGGAATCAGGTTGAACTGCTTCAGATGACT	300	
Db	241	GGGAATGATTTTGGTAAACGATCGAATCCACAGGAATCAGGTTGAACTGCTTCAGATGACT	300	
QY	301	TTCCGAGACGCTCCAGAGAATCTTCCGAGAATCATGCCCCAAGAGCCAGCAGAGGAAGAA	360	
Db	301	TTCCGAGACGCTCCAGAGAATCTTCCGAGAATCATGCCCCAAGAGCCAGCAGAGGAAGAA	360	
QY	361	AATGGTTTGAAGAAAGTCCACAGAGGATCTGGCCCCACAAATCATGGGAACAGCTGTGC	420	
Db	361	AATGGTTTGAAGAAAGTCCACAGAGGATCTGGCCCCACAAATCATGGGAACAGCTGTGC	420	
QY	421	CCCCGGGAAATCCAAAGTACCTTGGAGAAAGATTAACAAGACATCTGGACCCAAAAGGGGG	480	
Db	421	CCCCGGGAAATCCAAAGTACCTTGGAGAAAGATTAACAAGACATCTGGACCCAAAAGGGGG	480	
QY	481	AAACATGCTCGACCCACACAGACTGCGTGGAGAAAGCAGCTGGTGTATGAAGAGATC	540	
Db	481	AAACATGCTCGACCCACACAGACTGCGTGGAGAAAGCAGCTGGTGTATGAAGAGATC	540	
QY	541	ACGACCCCTGAGGAAGATGACGAGTAACCTCCCTCG	576	
Db	541	ACGACCCCTGAGGAAGATGACGAGTAACCTCCCTCG	576	

```

RESULT 4
US-10-117-937-599
; Sequence 599, Application US/10117937
; Publication No. US2003022039A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: Liu, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: CTLMM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07

```

```

; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 599
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-117-937-599

```

Query Match	100.0%	Score 576;	DB 15;	Length 576;
Best Local Similarity	100.0%;	Prod. No. 2.8e-174;		
Matches 576;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAACGGAGACGACGCGCTTTGCGAAGAGACCCGAGGATGATGCTCAAATATACGAGAAG	60	
DB	1	ATGAACGGAGACGACGCGCTTTGCGAAGAGACCCGAGGATGATGCTCAAATATACGAGAAG	60	
QY	61	TTACGAAGAAGCGCTTCGATGATATTGCCAAATACCTTCCTAAGAAGAAGTGGGAAAAGATG	120	
DB	61	TTACGAAGAAGCGCTTCGATGATATTGCCAAATACCTTCCTAAGAAGAAGTGGGAAAAGATG	120	
QY	121	AAATCCTCGAGAAAATCGCTTATGTGTATGAAGCTAAACTATGAGGTCATGACTATAA	180	
DB	121	AAATCCTCGAGAAAATCGCTTATGTGTATGAAGCTAAACTATGAGGTCATGACTATAA	180	
QY	181	CTAGGTTTCAAGGTCACCCCTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC	240	
DB	181	CTAGGTTTCAAGGTCACCCCTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC	240	
QY	241	GGGAATGATTTTGGTAAACGATCGAAACCCAGGAATCAGGTTGAAAGCTCCTCAGATGACT	300	
DB	241	GGGAATGATTTTGGTAAACGATCGAAACCCAGGAATCAGGTTGAAAGCTCCTCAGATGACT	300	
QY	301	TTCCGACAGCTCCAGAGAATCTTCCCGAAGATCATGCCAAGAACCCAGAGAGAGAA	360	
DB	301	TTCCGACAGCTCCAGAGAATCTTCCCGAAGATCATGCCAAGAACCCAGAGAGAGAA	360	
QY	361	AATGGTTTGAAGGAAGTGCCAGAGGATCTGGGCCACAAAATCATCGGAAACAGCTGTGC	420	
DB	361	AATGGTTTGAAGGAAGTGCCAGAGGATCTGGGCCACAAAATCATCGGAAACAGCTGTGC	420	
QY	421	CCCCGGGAAATCCAAAGTACTTTGGAGAGATTTAAACAAGACATCTGGACCCAAAAGGGGG	480	
DB	421	CCCCGGGAAATCCAAAGTACTTTGGAGAGATTTAAACAAGACATCTGGACCCAAAAGGGGG	480	
QY	481	AAACATGCTCGGACCCACAGACTGCGTGTAGAGAAACGACGCTGGTGGTTTATGAAGATC	540	
DB	481	AAACATGCTCGGACCCACAGACTGCGTGTAGAGAAACGACGCTGGTGGTTTATGAAGATC	540	
QY	541	AGCGACCTTGAGGAAGATGACGATTAACCTCCCTCG	576	
DB	541	AGCGACCTTGAGGAAGATGACGATTAACCTCCCTCG	576	

```

RESULT 5
US-09-849-602-13
; Sequence 13, Application US/09849602
; Publication No. US20030165834A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Chen, Yao-Tseng
; TITLE OF INVENTION: Colon Cancer Antigen Panel
; FILE REFERENCE: L0461/7105(JRV)
; CURRENT APPLICATION NUMBER: US/09/849,602
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 766
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-602-13

```

	Query Match	86.2%; Score 496.6; DB 10; Length 766;	
	Best Local Similarity	91.5%; Pred. No. 1e-148;	
	Matches 526; Conservative	0; Mismatches 49; Indels 0; Gaps 0;	
QY	1 ATGACCGAGACGCGCTTTGCAGAGAGACCAGGGATGATCTCAATATCAGAGAAG	60	
DB	92 ATGAACGAGACGCGCTTTGCAGAGAGACCAGGGATGATCTCAATATCAGAGAAG	151	
QY	61 TTACGAAAGGCCCTTCGATGATATTGCCAATACTTCTTAAGAAAGAGTGCGGAAGAAG	120	
DB	152 ATCCAAAAGGCCCTTCGATGATATTGCCAATACTTCTTAAGAAAGAGTGCGGAAGAAG	211	
QY	121 AAATCTCCGAGAAAAATCGTCTATGTGTATATCAAGCTAAAATATGAGGCTCATGACTAAA	180	
DB	212 AAAGCCTTCGAGAAAAATCTTCTATGTGTATATGAAGAAAGTATGAGGCTATGACTAAA	271	
QY	181 CTAGTTTTCAAGGTCACCCCTCCACACCTTTTCATGCGTAGTAAAAACGGGCTGCAGACTTCCCAC	240	
DB	272 CTAGTTTTCAAGGCCACCCCTCCCACTTCATGTGTATATTAACGGGCCGAGACTTCCAG	331	
QY	241 GGGAATGATTTTGGTAACGATCGAAACACAGGAATCAGTTGTAACGTCCTCAGATGACT	300	
DB	332 GGGAATGATTTTGGTAATATGACCTTAACCGTGGGAATCAGTTTGAACGTCCTCAGATGACT	391	
QY	301 TTCGGCAGCCCTCCAGAGAATCTTCCGGAAGATCATGCCAACAGACCCAGCAGAGGAAGA	360	
DB	392 TTCGGCAGGCTCCAGGGAATCTCCCGAAGATCATGCCAACAGACCCAGCAGAGGAAGA	451	
QY	361 AATGTTTGAAGAAATGCGCAGAGGATCTGCCCCACAGAGCATCTGCCCCACAAAATGATGGGAAGAGCTGTGC	420	
DB	452 AATGATTCGAGGAAGTGCAGAGGATCTGCCCCACAGAGCATCTGCCCCACAAAATGATGGGAAGAGCTGTGC	511	
QY	421 CCCCCGGGAAAATCCAAGTACCTTGGGAGAAGATTAAACAAGACATCTGGACCCCAAAGGGGG	480	
DB	512 CCCCCGGGAAAACCAACTACCTCTGAGAAGATTCAGAGAGATCTGGAGAGAGCTGTGC	571	
QY	481 AATGATTCGAGGAAGTGCAGAGGATCTGCCCCACAGAGCATCTGCCCCACAAAATGATGGGAAGAGCTGTGC	540	
DB	572 GAACATGCTCGGACCCACAGACTGCGTGCAGAGAAACAGCTGGTGGTATGAGAGATC	631	
QY	541 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC	575	
DB	632 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC	666	
RESULT 6			
US-10-207-655-84			
Sequence 84, Application US/10207655			
Publication No. US2003018592A1			
GENERAL INFORMATION:			
APPLICANT: Ledbetter, Jeffrey A.			
APPLICANT: Hayden-Jedbetter, Martha S.			
TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS			
FILE REFERENCE: 390069.401C1			
CURRENT APPLICATION NUMBER: US/10/207,655			
PRIORITY FILING DATE: 2002-07-25			
NUMBER OF SEQ ID NOS: 426			
SOFTWARE: PatentIn version 3.0			
SEQ ID NO 84			
LENGTH: 766			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-10-207-655-84			

QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 152 ATCCAAAGGCTTCGATGATATTCGCAAAATCTTCTTAAGAAAGAGTGGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGTCTATGACTAAA 180
Db 212 AAAGCCTCGGAGAAATCTTCTATGTATATGAAGAGAAAGTATGAGCTATGACTAAA 271
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTGTAGTAAACGGGCTCCAGACTTCCAC 240
Db 272 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTGTATTAACGGGCTCCAGACTTCCAG 331
QY 241 GGAATGATTTGGTAAAGATCGAAACCAACAGGAATCAGGTTGAACGCTCCTCAGATGACT 300
Db 332 GGAATGATTTGGTAAAGATCGAAACCAACAGGAATCAGGTTGAACGCTCCTCAGATGACT 391
QY 301 TTCGCGAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGAGAGAGAA 360
Db 392 TTCGCGAGCTCCAGGGAATCTCCCGAAGATCATGCCCAAGAGCCAGAGAGAGAA 451
QY 361 AATGTTTGAAGGAAGTGCAGAGGATCTGCGCCACAAATGATGGGAAACAGCTGTGC 420
Db 452 AATGATTCGAGGAAGTGCAGAGATCTGCGCCACAAATGATGGGAAAGAGCTGTGC 511
QY 421 CCCCCGGAATCCAAATACCTTTGAGAGAGATTAACAAGACATCTGGACCCCAAAAGGGG 480
Db 512 CCCCCGGAATCCAAATACCTTTGAGAGAGATTAACAAGACATCTGGACCCCAAAAGGGG 571
QY 481 AAACATGCTGACCCAGAGCTGCGTGAGAGAAAGAGCTGCTGTTTATGAAGATC 540
Db 572 GAACATGCTGACCCAGAGCTGCGTGAGAGAAAGAGCTGCTGTTTATGAAGATC 631
QY 541 AGCGACCTTGAGGAAGATGACGAGTAATCTCCCTC 575
Db 632 AGCGACCTTGAGGAAGATGACGAGTAATCTCCCTC 666

RESULT 8

US-09-833-039-2
; Sequence 2, Application US/09833039
; Publication No. US20030175960A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: US/09/833,039
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 09/409,455
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US 09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 2
; LENGTH: 931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-039-2

Query Match 86.2%; Score 496.6; DB 10; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.1e-148;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGAGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 99 ATGAACGAGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATCTTCTTAAGAAAGAGTGGGAAAGATG 120

Db 159 ATCCAAAGGCTTCGATGATATTCGCAAAATCTTCTTAAGAAAGAGTGGGAAAGATG 218
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGTCTATGACTAAA 180
Db 219 AAAGCCTCGGAGAAATCTTCTATGTATATGAAGAGAAAGTATGAGCTATGACTAAA 278
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTGTAGTAAACGGGCTCCAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTGTATTAACGGGCTCCAGACTTCCAG 338
QY 241 GGAATGATTTGGTAAAGATCGAAACCAACAGGAATCAGGTTGAACGCTCCTCAGATGACT 300
Db 339 GGAATGATTTGGTAAAGATCGAAACCAACAGGAATCAGGTTGAACGCTCCTCAGATGACT 398
QY 301 TTCGCGAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGAGAGAGAA 360
Db 399 TTCGCGAGCTCCAGGGAATCTCCCGAAGATCATGCCCAAGAGCCAGAGAGAGAA 458
QY 361 AATGTTTGAAGGAAGTGCAGAGGATCTGCGCCACAAATGATGGGAAACAGCTGTGC 420
Db 459 AATGATTCGAGGAAGTGCAGAGATCTGCGCCACAAATGATGGGAAAGAGCTGTGC 518
QY 421 CCCCCGGAATCCAAATACCTTTGAGAGAGATTAACAAGACATCTGGACCCCAAAAGGGG 480
Db 519 CCCCCGGAATCCAAATACCTTTGAGAGAGATTAACAAGACATCTGGACCCCAAAAGGGG 578
QY 481 AAACATGCTGACCCAGAGCTGCGTGAGAGAAAGAGCTGCTGTTTATGAAGATC 540
Db 579 GAACATGCTGACCCAGAGCTGCGTGAGAGAAAGAGCTGCTGTTTATGAAGATC 638
QY 541 AGCGACCTTGAGGAAGATGACGAGTAATCTCCCTC 575
Db 639 AGCGACCTTGAGGAAGATGACGAGTAATCTCCCTC 673

RESULT 9

US-10-177-277-2
; Sequence 2, Application US/10177277
; Publication No. US20030185844A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; APPLICANT: Ramesse, Hans Georg
; APPLICANT: Stvanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer In a Sample By Determi
; TITLE OF INVENTION: Expression of an SSX Gene, Peptides Derived From Said SSX Gene
; FILE REFERENCE: LUD 5556.1
; CURRENT APPLICATION NUMBER: US/10/177,277
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 132
; SEQ ID NO 2
; LENGTH: 931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-277-2

Query Match 86.2%; Score 496.6; DB 14; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.1e-148;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGAGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 99 ATGAACGAGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158

Qy	61	TTACGAAGGGCTTCGATGATATCCCAAATACTCTCTAAGAAAGATGGGAAAGATG	120
Db	159	ATCCAAAGGGCTTCGATGATATTCCAAATACTCTCTAAGCAAGATGGGAAAGATG	218
Qy	121	AAATCCTCGAGAAATCGTCTATCTGTATATGAAGCTAAACTATGAGTCACTATAA	180
Db	219	AAAGCCTCGAGAAATCTTCTATGTATATGAAGAGAGTATGAGCTATGACTATAA	278
Qy	181	CTAGGTTTCAAGGTCACCCCTCCCACTTTCATGTAGTAAACGGGCTGCAGACTTCCAC	240
Db	279	CTAGGTTTCAAGGCCACCCCTCCCACTTTCATGTGTAATAACCGGGCCGAAGACTTCCAG	338
Qy	241	GGGAATGATTTTGGTTAAACGATCGCAAAACCAAGGATCAGGTTGAACGTCCTCAGATGACT	300
Db	339	GGGAATGATTTGGATATATGACCTTAACCGTGGAAATCAGGTTGAACGTCCTCAGATGACT	398
Qy	301	TTCCGAGCCTCCAGAGAAATCTTCCGAAGATCATGCCCAAGAACCGACGAGAGAAAGAA	360
Db	399	TTCCGAGGCTTCCAGGGAATCTCCCGAAGATCATGCCCAAGAACCGACGAGAGAAAGAA	458
Qy	361	AATGGTTTGAAGGAAGTCCACAGAGSCATCTGGCCCAACAAAATGATGGGAAACAGCTGTGC	420
Db	459	AATGATTCGAGGAAGTCCACAGAGCATCTGGCCCAACAAAATGATGGGAAAGAGCTGTGC	518
Qy	421	CCCCCGGAAATCCAAGTACCTTGAGAGAGATTAAACAAGACATCTGGACCCCAAAAGGGGG	480
Db	519	CCCCCGGAAACCAACTACCTCTGAGAAGATTACAGAGAGATCTGGACCCCAAAAGGGGG	578
Qy	481	AAACATGCTTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGGTTATGAGAGATC	540
Db	579	GAACATGCTTGGACCCACAGACTCGGTGAGAGAAACAGCTGGTGGTTATGAGAGATC	638
Qy	541	AGGACCCCTGAGGAAGATGACGAGTAATCCCCCTC	575
Db	639	AGGACCCCTGAGGAAGATGACGAGTAATCCCCCTC	673

Qy 541 AGCGACCTCGAGGAAGATGACGAGTAACGCCCTC 575
|||||
Db 639 AGCGACCTCGAGGAAGATGACGAGTAACGCCCTC 673

US-09-954-531-1392

```

Query Match      86.2%; Score 496.6; DB 9; Length 1309;
Best Local Similarity 91.5%; Pred. No. 1.4e-148;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGAGCGCCCTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG 60
DB 99 ATGAACGGAGAGCGCCCTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG 158
QY 61 TTACGAAAGGCGCTTCGATGATATTCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
DB 159 ATCCAAAGGCGCTTCGATGATATTCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 218
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCTATGACTAAA 180
DB 219 AAAGCTCGGAGAAATCTTCTATGTATATGAAGCTAAACTATGAGGTCTATGACTAAA 278
QY 181 CTAGGTTTCAAGGTACCCCTCCACCTTTCATGCGTAGTAAACGGGCTCGAGACTTCCAC 240
DB 279 CTAGGTTTCAAGGCGCCCTTCACACCTTTCATGTGTATATGAAGCTAAACTATGACTAAA 338
QY 241 GGAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACTCTCTCAGATGACT 300
DB 339 GGAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACTCTCTCAGATGACT 398
QY 301 TTCCGAGGCTCCAGAGATCTTCCCGAAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
DB 399 TTCCGAGGCTCCAGAGATCTTCCCGAAGATCATGCCCAAGAGCCAGCAGAGGAGAA 458
QY 361 AATGTTTGAAGGAGTGCAGAGGATCTGGCCCAAAATCATGGGAAACAGCTGTGC 420
DB 459 AATGTTTGAAGGAGTGCAGAGGATCTGGCCCAAAATCATGGGAAACAGCTGTGC 518
QY 421 CCCCCGGGAAATCCAACTACTCTGAGAGATTAACAAGACATCTGCAACCCAAAGGGGG 480
DB 519 CCCCCGGGAAATCCAACTACTCTGAGAGATTAACAAGACATCTGCAACCCAAAGGGGG 578
QY 481 AAACATGCTGGAGCCACAGATCTGCGTGAAGAGAGAGAGAGAGAGAGATC 540
DB 579 GAACATGCTGGAGCCACAGATCTGCGTGAAGAGAGAGAGAGAGAGATC 638
QY 541 AGCAGCCCTGAGGAGATGACGAGTACTCCCTC 575
DB 639 AGCAGCCCTGAGGAGATGACGAGTACTCCCTC 673

RESULT 12
US-10-159-563-402
; Sequence 402, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS
; FILE REFERENCE: 11613.56US11
; CURRENT FILING DATE: 2002-12-09
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 402
; LENGTH: 1249
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-402
Query Match      86.1%; Score 496; DB 15; Length 1249;
Best Local Similarity 91.3%; Pred. No. 2.1e-148;

```

```

Matches 526; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 ATGAACGGAGAGCGCCCTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG 60
DB 54 ATGAACGGAGAGCGCCCTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG 113
QY 61 TTACGAAAGGCGCTTCGATGATATTCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
DB 114 ATCAAAAGGCGCTTCGATGATATTCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 173
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCTATGACTAAA 180
DB 174 AAAGCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCTATGACTAAA 233
QY 181 CTAGGTTTCAAGGTACCCCTCCACCTTTCATGCGTAGTAAACGGGCTCGAGACTTCCAC 240
DB 234 CTAGGTTTCAAGGCGCCCTTCACACCTTTCATGCGTAGTAAACGGGCTCGAGACTTCCAC 293
QY 241 GGAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACTCTCTCAGATGACT 300
DB 294 GGAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACTCTCTCAGATGACT 353
QY 301 TTCCGAGGCTCCAGAGATCTTCCCGAAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
DB 354 TTCCGAGGCTCCAGAGATCTTCCCGAAGATCATGCCCAAGAGCCAGCAGAGGAGAA 413
QY 361 AATGTTTGAAGGAGTGCAGAGGATCTGGCCCAAAATCATGGGAAACAGCTGTGC 420
DB 414 AATGTTTGAAGGAGTGCAGAGGATCTGGCCCAAAATCATGGGAAACAGCTGTGC 473
QY 421 CCCCCGGGAAATCCAACTACTCTGAGAGATTAACAAGACATCTGCAACCCAAAGGGGG 480
DB 474 CCCCCGGGAAATCCAACTACTCTGAGAGATTAACAAGACATCTGCAACCCAAAGGGGG 533
QY 481 AAACATGCTGGAGCCACAGATCTGCGTGAAGAGAGAGAGAGAGATC 540
DB 534 GAACATGCTGGAGCCACAGATCTGCGTGAAGAGAGAGAGAGATC 593
QY 541 AGCAGCCCTGAGGAGATGACGAGTAACTCCCTC 576
DB 594 AGCAGCTCTGAGAGATGATGAGTAACTCCCTC 629

RESULT 13
US-09-975-856-2
; Sequence 2, Application US/09975856
; Publication No. US20030023057A1
; GENERAL INFORMATION:
; APPLICANT: Gure, Ali O.; Tureci, Ozlem; Sahin, Ugur; Tsang, Solan;
; Scanlan, Matthew J.; Knuth Alexander; Pfreundschuh, Michael; Old,
; Lloyd J.; Chen, Yao-Tseng
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding SSX
; Family
; Members And Uses Thereof
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/975,856
; FILING DATE: 11-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/796,780
; FILING DATE: 2001-03-01

```

ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. US20030023057Alman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5480
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3168
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-975-856-2

Query Match 83.7%; Score 482.2; DB 10; Length 576;
Best Local Similarity 89.9%; Pred. No. 3.7e-144;
Matches 517; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 1 ATGAACGGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60

QY 61 TTACGAAGCGCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
Db 61 ATGCAAAAGCGCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120

QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTTCATGACTAAA 180
Db 121 AAAGCCTCGGAGAAATCATCTATGTATATGAAGAAAGTATGAGGCCATGACTAAA 180

QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGGCGACCCCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAG 240

QY 241 GCGAATGATTTGGTAAACGATCGAACCACAGGAATCAGGTTGAACGTCTCAGATGACT 300
Db 241 GCGAATGATTTGAATGATACCTTAACCGTGGGAATCAGGTTGAACATCTCAGATGACT 300

QY 301 TTCGGCAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
Db 301 TTCGGCAGCCTCCAGGAAATCTTCCCGAAGATCATGCCCGAGAGCCAGCAGAGGAAGAA 360

QY 361 AATGTTTGAAGAAAGTGCAGAGGATCTGGCCCAAAATATGATGGGAAACAGCTGTGC 420
Db 361 AATGATTCAAAGGAGTGCAGAGGATCTGGCCCAAGAACTATGGGAAACAGCTGTGC 420

QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCAAAGGGGG 480
Db 421 CCTCAGGAAACTTAATACCTCTGAGAGGTTTAAAGACATCTGGACCCAAAGGGGG 480

QY 481 AAACATGCTTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 481 AAACATGCTTGGACCCACAGAGTGGTGAGAGAAAGCAACTGGTGGATTATGAAGATC 540

QY 541 AGCGACCTCAGGAAGATGACGAGTAACCTCCCTC 575
Db 541 AGCGACCTCAGGAAGATGACGAGTAACCTCCCTC 575

RESULT 14
US-09-833-039-6
; Sequence 6, Application US/09833039
; Publication No. US20030175960A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfrendschuh, Michael
; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: US/09/833,039
; CURRENT FILING DATE: 2001-04-12

PRIOR APPLICATION NUMBER: US 09/409,455
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: US 09/344,040
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,839
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: US 08/851,130
PRIOR FILING DATE: 1997-05-05
NUMBER OF SEQ ID NOS: 129
SEQ ID NO 6
LENGTH: 576
TYPE: DNA
ORGANISM: Homo sapiens
US-09-833-039-6

Query Match 83.7%; Score 482.2; DB 10; Length 576;
Best Local Similarity 89.9%; Pred. No. 3.7e-144;
Matches 517; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 1 ATGAACGGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60

QY 61 TTACGAAGCGCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
Db 61 ATGCAAAAGCGCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120

QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTTCATGACTAAA 180
Db 121 AAAGCCTCGGAGAAATCATCTATGTATATGAAGAAAGTATGAGGCCATGACTAAA 180

QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGGCGACCCCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAG 240

QY 241 GCGAATGATTTGGTAAACGATCGAACCACAGGAATCAGGTTGAACGTCTCAGATGACT 300
Db 241 GCGAATGATTTGAATGATACCTTAACCGTGGGAATCAGGTTGAACATCTCAGATGACT 300

QY 301 TTCGGCAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
Db 301 TTCGGCAGCCTCCAGGAAATCTTCCCGAAGATCATGCCCGAGAGCCAGCAGAGGAAGAA 360

QY 361 AATGTTTGAAGAAAGTGCAGAGGATCTGGCCCAAAATATGATGGGAAACAGCTGTGC 420
Db 361 AATGATTCAAAGGAGTGCAGAGGATCTGGCCCAAGAACTATGGGAAACAGCTGTGC 420

QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCAAAGGGGG 480
Db 421 CCTCAGGAAACTTAATACCTCTGAGAGGTTTAAAGACATCTGGACCCAAAGGGGG 480

QY 481 AAACATGCTTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 481 AAACATGCTTGGACCCACAGAGTGGTGAGAGAAAGCAACTGGTGGATTATGAAGATC 540

QY 541 AGCGACCTCAGGAAGATGACGAGTAACCTCCCTC 575
Db 541 AGCGACCTCAGGAAGATGACGAGTAACCTCCCTC 575

RESULT 15
US-10-177-277-6
; Sequence 6, Application US/10177277
; Publication No. US20030185844A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfrendschuh, Michael
; APPLICANT: Ramensee, Hans Georg
; APPLICANT: Stewanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determ:
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: Expression of an Ssx Gene, Peptides Derived from Said Ssx Gene
; CURRENT FILING DATE: 2001-04-12


```
/ FILE REFERENCE: LUD 5556.1
/ CURRENT APPLICATION NUMBER: US/10/177,277
/ CURRENT FILING DATE: 2002-06-21
/ PRIOR APPLICATION NUMBER: US/09/344,040
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: US 09/105,839
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: US 08/851,130
/ PRIOR FILING DATE: 1997-05-05
/ NUMBER OF SEQ ID NOS: 132
/ SEQ ID NO 6
/ LENGTH: 576
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-177-277-6

Query Match      83.7%; Score 482.2; DB 14; Length 576;
Best Local Similarity 89.9%; Pred. No. 3.7e-144;
Matches 517; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY      1 ATGAACGGAGACACGCCCTTTGCAAGGAGACCCAGGGATGATGCTCAAATATCAGAGAAG 60
Db      |||
QY      1 ATGAACGGAGACACGCCCTTTGACGAGACCTAGGGTTGGTTCTCAATATCCACAGAAG 60
Db      |||
QY      61 TTACGAAAGGCCTTCGATGATATTGCCAAATCTTCTTAAGAAAGAGTGGGAAAAGATG 120
Db      |||
QY      61 ATGCAAAAGGCCTTCGATGATATTGCCAAATCTTCTCTGAGAAAGAGTGGGAAAAGATG 120
Db      |||
QY      121 AAATCTCCGAGAAATCTCTATGTATATGAAGCTAAACTATGAGTCTGACTATAA 180
Db      |||
QY      121 AAGGCTCCGAGAAATCTCTATGTATATGAAGAAAGTATGAGGCCATGACTATAA 180
Db      |||
QY      181 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db      |||
QY      181 CTAGGTTTCAAGGCCACCTCCACCTTTCATGCGTAAATACGGGTCCGACACTTCCAG 240
Db      |||
QY      241 GGGATGATTTTGGTAAAGATCGAAACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Db      |||
QY      241 GGGATGATTTTGAATGATACCTTAACCGTGGGAATCAGGTTGAACATCCTCAGATGACT 300
Db      |||
QY      301 TTCGGCAGCCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCACAGAGAGAGAA 360
Db      |||
QY      301 TTCGGCAGCCTCCAGGAATCTTCCGAGAGATCAOCCCCGAGAGCCACAGAGAGAGGA 360
Db      |||
QY      361 AATGTTTGAAGGAAGTGCAGAGGATCTGGCCCCACAAAATGATGGGAAAACAGCTGTGC 420
Db      |||
QY      361 AATGATTCAAAGGAGTGCAGAGATCTGGCCCCACAGAACTGGGAAAACAGCTGTGC 420
Db      |||
QY      421 CCCCAGGGAATCCAGTACCTTGGAGAGATTAAACAGATCTGGACCCCAAAAGGGGG 480
Db      |||
QY      421 CCTCAGGAAAACCTAAATACCTCTGAGAAAGTTAACAGACATCTGGACCCCAAAAGGGGG 480
Db      |||
QY      481 AAACATGCTGGACCCACAGATCTGGTGAAGAAAGCAGCTGGTGTATGAAGATC 540
Db      |||
QY      481 AAACATGCTGGACCCACAGATCTGGTGAAGAAAGCAGCTGGTGTATGAAGATC 540
Db      |||
QY      541 AGCGACCTTGAGGAAGATGACGAGTAATCCCTC 575
Db      |||
QY      541 AGCGACCTTGAGGAAGATGACGAGTAATCCCTC 575
Db      |||
```

Search completed: March 31, 2004, 18:32:55
Job time : 402 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 31, 2004, 13:39:56 ; Search time 14 Seconds
(without alignments)
4284.630 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 1025

Sequence: 1 ATGACGAGGAGACGACCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/cgn2_1/USPTO spoel_p/US09975856/runat 31032004 132903 14738/app query.fasta 1.775
-DB=SwissProt 42 -Qfmt=fastan -SUFFIX=xsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LISF=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09975856 @CGN_1 1 16 @runat 31032004 132903 14738 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	ID	Description
1	1004	98.0	188	1	SSX4_HUMAN	O60224 homo sapien
2	801	78.1	188	1	SSX3_HUMAN	Q99909 homo sapien
3	798	77.9	188	1	SSX2_HUMAN	Q16385 homo sapien
4	784	76.5	188	1	SSX1_HUMAN	Q16384 homo sapien
5	780	76.1	188	1	SSX5_HUMAN	O60225 homo sapien
6	113.5	11.1	289	1	ZN75_HUMAN	P51815 homo sapien
7	101	9.9	642	1	ZN98_HUMAN	Q8cd17 homo sapien
8	94	9.2	595	1	ZN17_HUMAN	O96p66 homo sapien
9	89.5	8.9	1505	1	PK37_RAT	Q07073 rattus norv
10	88.5	8.6	531	1	FM01_CANFA	Q951a2 canis famul
11	88	8.6	743	1	REP1_MOUSE	O54916 mus musculu
12	87.5	8.7	531	1	FM01_PIG	P16549 sus scrofa
13	87	8.5	744	1	REP1_HUMAN	Q46d71 homo sapien
14	86.5	8.4	396	1	SK11_CHICK	P98435 gallus gall
15	86.5	8.4	446	1	ZN38_HUMAN	P17036 homo sapien
16	86.5	8.4	488	1	ZF92_MOUSE	Q62396 mus musculu
17	86.5	8.4	572	1	TC17_MOUSE	Q61751 mus musculu
18	86	8.4	480	1	WR61_ARATH	Q8vvv6 arabidopsis

19	86	8.4	834	1	CASL_HUMAN	Q14511 homo sapien
20	85.5	8.3	680	1	Z334_HUMAN	Q9hcz1 homo sapien
21	85.5	8.3	738	1	ZN84_HUMAN	P51323 homo sapien
22	84.5	8.2	429	1	HISX_METJA	Q5851 methanococc
23	84.5	8.2	463	1	Z331_HUMAN	Q9ngx6 homo sapien
24	84.5	8.4	534	1	FM01_RABIT	P17636 oryctolagus
25	84	8.2	524	1	Z479_HUMAN	Q96jca4 homo sapien
26	83.5	8.1	636	1	ZF90_MOUSE	Q15677 mus musculu
27	83.5	8.1	1102	1	ADNP_HUMAN	Q9h2p0 homo sapien
28	83.5	8.1	1630	1	MSP1_PLAFA	P04332 plasmodium
29	83.5	8.1	1639	1	MSF1_PLAFA	P04333 plasmodium
30	82.5	8.0	553	1	Z334_HUMAN	O75467 homo sapien
31	82.5	8.0	643	1	ZN74_HUMAN	Q16587 homo sapien
32	82.5	8.0	1035	1	RRPO_BMYVF	P09507 beet wester
33	81.5	8.0	504	1	Z205_HUMAN	O95201 homo sapien
34	81.5	8.0	639	1	GLOX_SOYBN	P11827 glycine max
35	81.5	8.0	1395	1	SP41_YEAST	P38904 saccharomyc
36	81.5	8.0	1443	1	DPO3_MYCPN	P75080 mycoplasma
37	81	7.9	458	1	ZN19_HUMAN	P17023 homo sapien
38	81	7.9	751	1	Z337_HUMAN	Q9y3m9 homo sapien
39	81	8.0	933	1	PERT_CANFA	Q8hyb7 canis famul
40	80.5	7.9	439	1	ZN10_HUMAN	P21506 homo sapien
41	80.5	7.9	470	1	Z436_HUMAN	Q9c0f3 homo sapien
42	80.5	7.9	659	1	Z304_HUMAN	Q9hcz3 homo sapien
43	80.5	7.9	998	1	GTFI_HUMAN	P78347 h general t
44	79.5	7.9	531	1	FM01_HUMAN	Q01740 homo sapien
45	79.5	7.8	626	1	Z471_HUMAN	Q9bx82 homo sapien

ALIGNMENTS

RESULT 1	SSX4_HUMAN	STANDARD;	PRT;	188 AA.
ID	SSX4_HUMAN	Q9UJY9;		
AC	O60224; Q9UJY9;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	SSX4 protein.			
GN	SSX4.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98021352; PubMed=9378559;			
RA	Gure A.O., Tuercio O., Sahin U., Tsang S., Scanlan M.J., Jager E.,			
RA	Knuth A., Pfeundschn M., Old L.J., Chen Y.-T.;			
RT	"SSX: a multigene family with several members transcribed in normal			
RT	testis and human cancer."			
RL	Int. J. Cancer 72:965-971(1997).			
RP	SEQUENCE FROM N.A.			
RN	[2]			
RC	TISSUE=Urinary bladder;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Stapchenko L., Varusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,			
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;			

"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

[3]
SEQUENCE OF 1-155 FROM N.A.
Bleichschmidt K., Nyakatura G., Strom T.M., Drescher B., Menzel U., Meindl A., Rosenthal A.;
Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: Could act as a modulator of transcription.
-!- SIMILARITY: Belongs to the SSX family.
-!- SIMILARITY: Contains 1 KRAB-related domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U90841; AAC05820.1; -;
EMBL; BC005325; AAH05325.1; -;
EMBL; AF196972; AAF06796.1; -;
Genew; HGNC:11338; SSX4.
MIM; 300326;
InterPro; IPR001909; KRAB.
InterPro; IPR003655; KRAB-related.
Pfam; PF01352; KRAB; 1.
SMART; SM00349; KRAB; 1.
PROSITE; PS50806; KRAB RELATED; 1.
Multigene family; Transcription regulation.
FT DOMAIN 20 83 KRAB-RELATED.
SQ SEQUENCE 188 AA; 21858 MW; 4ACA2A8737507AE5 CRC64;

Alignment Scores:
Pred. No.: 1-37e-91 Length: 188
Score: 1004.00 Matches: 188
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.95% Indels: 0
DB: 1 Gaps: 0

US-09-975-856-1 (1-576) x SSX4_HUMAN (1-188)

QY 1 ATGAACGAGACGACCCCTTTCAGAGAGACCCAGGATGATCTCAATATCAGAGAG 60
DB 1 MetAsnGlyAspAspAlaPheAlaArgProArgAspAlaGlnIleSerGluLys 20
QY 61 TTACGAAAGCGCTTCGATGATATTCGCAATCTCTCTAAGAAAGAGTGGGAAAGATG 120
DB 21 LeuArgLysAlaPheAspAspIleAlaLysTyrPheSerLysLysGluTrpGluLysMet 40
QY 121 AAATCTCCGAGAAATCGTCTATGTGTATATGAAGCTAACTATCAGGTCTGACTAA 180
DB 41 LysSerSerGluLysIleValTyrValTyrMetLysLeuAsnTyrGluValMetThrLys 60
QY 181 CTAGTTTTCAGGTTCACCTCCACCTTCATCGTAGTAAGCGGCTGCAGACTTCCAC 240
DB 61 LeuGlyPheLysValThrLeuProPheMetArgSerLysArgAlaAlaAspPheHis 80
QY 241 GGAAGTATTGGTGAACCATGACCAACACAGCAATCAGTTGACGCTCTCAGTACT 300
DB 81 GlyAsnAspPheGlyAsnAspArgAsnHisArgAsnGlnValGluArgProGlnMetThr 100
QY 301 TTCGCGAGCCTCCAGAGATCTTCCGAGATCATGCCAAGAGCCAGCAGAGAGAA 360
DB 101 PheGlySerLeuGlnArgIlePheProLysIleMetProLysLysProAlaGluGlu 120
QY 361 ATGTTTGAAGAGTGCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
DB 121 AsnGlyLeuLysGluValProGluAlaSerGlyProGlnAsnAspGlyLysGlnLeuCys 140
QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAAACAGACATCTCGACCCCAAGGGG 480

Db 141 ProProGlyAsnProSerThrLeuGluLysIleAsnLysThrSerGlyProLysArgGly 160
QY 481 AAACATGCTGAGCCACAGACTGCTGAGAGAAAGCAGCTGCTGTTTATGAGAGATC 540
DB 161 LysHisAlaTrpThrHisArgLeuArgGluArgLysGlnLeuValTyrGluGluIle 180
QY 541 AGCAGCCCTGAGCAAGATGACGAG 564
DB 181 SerAspProGluGluAspAspGlu 188

RESULT 2

SSX3_HUMAN
ID SSX3_HUMAN STANDARD; PRT; 188 AA.
AC Q99509; O60223;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE SSX3 protein.
GN SSX3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibrosarcoma, and Testis;
RX MEDLINE=96302330; PubMed=8697803;
RA de Leeuw B., Balenans M., Geurts van Kessel A.;
RT "A novel Kruppel-associated box containing the SSX gene (SSX3) on the human X chromosome is not implicated in t(X;18)-positive synovial sarcomas";
RT Cytogenet. Cell Genet. 73:179-183 (1996).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98021352; PubMed=9378559;
RA Gure A.O., Tuereci O., Sahin U., Tsang S., Scanlan M.J., Jager E., Knuth A., Pfundschnuh M., Old L.J., Chen Y.-T.;
RT "SSX: a multigene family with several members transcribed in normal testis and human cancer";
RL Int. J. Cancer 72:965-971 (1997).
CC -!- FUNCTION: Could act as a modulator of transcription.
CC -!- SIMILARITY: Belongs to the SSX family.
CC -!- SIMILARITY: Contains 1 KRAB-related domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; S82471; AAB37436.2; -;
EMBL; U90840; AAC05819.1; -;
Genew; HGNC:11337; SSX3.
MIM; 300325; -;
InterPro; IPR001909; KRAB.
InterPro; IPR003655; KRAB-related.
Pfam; PF01352; KRAB; 1.
SMART; SM00349; KRAB; 1.
PROSITE; PS50806; KRAB-RELATED; 1.
Multigene family; Transcription regulation.
FT DOMAIN 20 83 KRAB-RELATED.
FT CONFLICT 95 95 L -> Q (IN REF. 2).
SQ SEQUENCE 188 AA; 21682 MW; 368C7487C1450E72 CRC64;

Alignment Scores:
Pred. No.: 1.85e-71 Length: 188
Score: 801.00 Matches: 152
Percent Similarity: 84.57% Conservative: 7
Best Local Similarity: 80.85% Mismatches: 29
Query Match: 78.15% Indels: 0

```

DB:      1      1      0
US-09-975-856-1 (1-576) x SSX2_HUMAN (1-188)

QY      1  ATGACGGAGAGCAGCGCTTTTCAGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Dbb      1  MetAsnGlyAspThrPheAlaArgArgProThrValGlyAlaGlnIleProGluLys 20
QY      61  TTACGAAAGCGCTTCGATGATATGCAATATCTCTTCAAGAAAGAGTGGGAAAGATG 120
Dbb      21  IleGlnLysAlaPheAspAlaLysTyrPheSerLysGluGluTrpGluLysMet 40
QY      121  AAATCCTCGGAGAAATCTCTATGATATGAAGTAAATATCAGATCATGACTAAA 180
Dbb      41  LysValSerGluLysIleValTyrMetLysArgLysTyrGluAlaMetThrLys 60
QY      181  CTAGCTTTCAAGGTCACCTCCACCTTCATGCGCTAGTAAACGGGCTGCACATCCAC 240
Dbb      61  LeuGlyPheLysAlaIleLeuProSerPheMetArgAsnLysArgValThrAspPheGln 80
QY      241  GGAATGATATTTGGTAACGATCGAAACACACAGGAATCAGTTGAACTGCTCAGATGACT 300
Dbb      81  GlyAsnAspPheAspAsnAspProAsnArgGlyAsnGlnValLeuArgProGlnMetThr 100
QY      301  TTCGCAGCGCTCCAGAGAAATCTCCGAGAGATCATGCCCCAAGAACCGACGAGAGAA 360
Dbb      101  PheGlyArgLeuGlnGlyIlePheProLysIleMetProLysLysProAlaGluGlu 120
QY      361  AATGTTTGAAGGAAGTGCACAGCATCTGCCCCCAAAATGATGGGAACAGCTGTC 420
Dbb      121  AsnValSerLysGluValProGluAlaSerGlyProGlnAspGlyLysGlnLeuCys 140
QY      421  CCCCAGGAAATCCAGTACCTTGGAGAGATTAACAAGACATCTCGACCCCAAAAGGGGG 480
Dbb      141  ProProGlyLysProThrThrSerGluLysIleAsnMetIleSerGlyProLysArgGly 160
QY      481  AAACATGCTGACCCACAGCATCTGGTGGAGAAAGAGCTGGTGGTTATGAAGATC 540
Dbb      161  GluHisAlaTrpThrHisArgLeuArgGluArgLysGlnLeuValIleTyrGluGlu 180
QY      541  AGCGACCTCTGAGGAGAGATGACGAG 564
Dbb      181  SerAspProGluGluAspGlu 188

RESULT 3
SSX2_HUMAN
ID      SSX2_HUMAN      STANDARD;      PRT;      188 AA.
AC      Q16385; Q16404; Q96IP7;
DT      15-DEC-1998 (Rel. 37, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      SSX2 protein (Synovial sarcoma, X breakpoint 2) (SSX) (HOM-MEL-40).
GN      SSX2.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Testis;
EX      MEDLINE=92292374; PubMed=7539744;
RA      Crew A.J., Clark J., Fisher C., Gill S., Grimer R., Chand A.,
RA      Shipley J., Gusterson B.A., Cooper C.S.;
RT      "Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with
RT      homology to the Kruppel-associated box in human synovial sarcoma.";
RL      EMBO J. 14:2333-2340(1995).
RP      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Skin;
RC      MEDLINE=22388257; PubMed=12477932;
RC      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

```

```

RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., McQuellano N.A., Peters G.J., Toshiyuki S., Carninci P., Prange C.,
RA      Bosak S.A., LeGuen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
RT      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RL      [3]
RN      [3]
RP      SEQUENCE OF 111-188 FROM N.A.
RC      TISSUE=Synovial sarcoma;
RX      MEDLINE=95384157; PubMed=7655467;
RA      de Leeuw B., Bailemans M., Olde Weghuis D., Geurts van Kessel A.;
RT      "Identification of two alternative fusion genes, SYT-SSX1 and
RT      SYT-SSX2, in t(X;18)(p11.2;q11.2)-positive synovial sarcomas.";
RL      Hum. Mol. Genet. 4:1097-1099(1995).
RN      [4]
RP      SEQUENCE OF 111-188 FROM N.A. (SSX1-SSX2 FUSION PROTEIN).
RC      TISSUE=Synovial sarcoma;
RX      MEDLINE=95038836; PubMed=7951320;
RA      Clark J., Roques P.J., Crew A.J., Gill S., Shipley J., Chan A.M.-L.,
RA      Gusterson B.A., Cooper C.S.;
RT      "Identification of novel genes, SYT and SSX, involved in the
RT      t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma.";
RL      Nat. Genet. 7:502-508(1994).
RN      [5]
RP      SEQUENCE OF 68-116 FROM N.A. (SSX1-SSX2 FUSION PROTEIN).
RX      MEDLINE=96094743; PubMed=7495284;
RA      Fligman I., Leonardo F., Jhanwar S.C., Gerald W.L., Woodruff J.,
RA      Ladanyi M.;
RT      "Molecular diagnosis of synovial sarcoma and characterization of a
RT      variant SYT-SSX2 fusion transcript.";
RL      Am. J. Pathol. 147:1592-1599(1995).
CC      -!- FUNCTION: Could act as a modulator of transcription.
CC      -!- TISSUE SPECIFICITY: Expressed at high level in the testis.
CC      Expressed at low level in thyroid. Not detected in tonsil, colon,
CC      lung, spleen, prostate, kidney, striated and smooth muscles.
CC      Detected in rhabdomyosarcoma and fibrosarcoma cell lines. Not
CC      detected in mesenchymal and epithelial cell lines.
CC      -!- DISEASE: Participates in a t(X;18)(p11.2;q11.2) chromosomal
CC      translocation that is specifically found in more than 80% of
CC      synovial sarcoma and produces the SSX1-SSX1 or SSX1-SSX2 fusion
CC      products. These hybrid proteins are probably responsible for
CC      transforming activity. Heterogeneity in the position of the
CC      breakpoint can occur (low frequency).
CC      -!- SIMILARITY: Belongs to the SSX family.
CC      -!- SIMILARITY: Contains 1 KRAB-related domain.
CC      [1]
CC      [2]
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; X86175; CAA60111.1; --
CC      EMBL; BC007343; AAH07343.1; --
CC      EMBL; BC016957; AAH16957.1; --
CC      EMBL; S79332; AAB35379.1; --
CC      EMBL; X79200; -- NOT ANNOTATED CDS.
CC      EMBL; S79894; AAB35674.1; ALT_INIT.
CC      PIR; S55058; S55058.
CC      Genew; HGNC:11336; SSX2.

```

```

DR MIM; 300192; -.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PSS0806; KRAB RELATED; 1.
KW Chromosomal translocation; Proto-oncogene; Multigene family;
KW Transcription regulation.
FT DOMAIN 20 83
FT SITE 68 69
FT SITE 110 111
FT SITE 169 169
FT CONFLICT 169 169
FT CONFLICT 188 AA; 21620 MW; BF5D18AA5F45B1B1 CRC64;
SQ SEQUENCE 188 AA; 21620 MW; BF5D18AA5F45B1B1 CRC64;

Alignment Scores:
Pred. No.: 3,66e-71 Length: 188
Score: 798.00 Matches: 150
Percent Similarity: 86.17% Conservative: 12
Best Local Similarity: 79.79% Mismatches: 26
Query Match: 77.85% Indels: 0
DB: Gaps: 1

US-09-975-856-1 (1-576) x SSX2 HUMAN (1-188)
QY 1 ATGACGGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAATATCAGAGAAG 60
Db 1 MetAsnGlyAspAlaPheAlaArgProThrValGlyAlaGlnIleProGluLys 20
QY 61 TTACGAAAGCCCTTCGATGATATGCAATATCTCTTAAGAAAGATGGGAAAAGATG 120
Db 1 IleGlnLysAlaPheAspAlaLysTyrPheSerLysGluLysMet 40
QY 121 AAATCTCGAGAAATCGCTATGTTATATCAAGCTAACTATCAGGTCATGACTAA 180
Db 1 LysAlaSerGluLysIlePheTyrValTyrMetLysArgLysTyrGluAlaMetThrLys 60
QY 181 CTAGTTTCAAGTCCACCTCCACCTTTTCATCGTAGTAAGCGGCTGCAGACTTCCAC 240
Db 1 LeuGlyPheLysAlaThrLeuProPheMetCysAsnLysArgAlaGluAspPheGln 80
QY 241 GGAATCAITTTGGTACGATCAACACCAACCAAGAAATCAGTTGAAGTCTCAGATGACT 300
Db 1 GlyAsnAspLeuAspAsnAspProAsnArgGlyAsnGlnValGluArgProGlnMetThr 100
QY 301 TTCCGAGCCTCCAGAGAATCTTCCGAGATCATGCCAAGATCATGCCAAGACCAAGAGAA 360
Db 1 PheGlyArgLeuGlnGlyIleSerProLysIleMetProLysLysProAlaGluGly 120
QY 361 AATGTTTGAAGAGTGCAGAGGATCTGCCCCAATAATGATGGAAACACAGCTGTGC 420
Db 1 AsnAspSerGluGluValProGluAlaSerGlyProGlnAsnAspGlyLysGluLeuCys 140
QY 421 CCCCCGGAAATCCAGTACCTTGGAGAGATTAACAAGACATCTGGACCCAAAGGGG 480
Db 1 ProProGlyLysProThrThrSerGluLysIleHisGluArgSerGlyProLysArgGly 160
QY 481 AAACATGCTGACCCACACACTGCTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 1 GluHisAlaThrThrHisArgLeuArgGluArgLysGlnLeuValIleTyrGluGluIle 180
QY 541 AGGACCTCGAGAGATGACGAG 564
Db 181 SerAspProGluGluAspAspGlu 188

RESULT 4
SSX1_HUMAN
ID SSX1_HUMAN STANDARD; PRT; 188 AA.
AC Q16384;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)

```

```

DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SSX1 protein (Synovial sarcoma, X breakpoint 1).
GN OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibrosarcoma;
RX MEDLINE=95292974; PubMed=7539744;
RA Crew A.J., Clark J., Fisher C., Gall S., Grimer R., Chand A.,
RA Shipley J., Gusterson B.A., Cooper C.S.;
RT "Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with
RT homology to the Kruppel-associated box in human synovial sarcoma.";
RL EMBO J. 14:2333-2340(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Locantillo N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RN [3]
RP Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 111-188 FROM N.A.
RC TISSUE=Synovial sarcoma;
RX MEDLINE=95384157; PubMed=7655467;
RA de Leeuw B., Balemans M., Olde Weghuis D., Geurts van Kessel A.;
RT "Identification of two alternative fusion genes, SYT-SSX1 and
RT SYT-SSX2, in t(X;18)(p11.2;q11.2)-positive synovial sarcomas.";
RL Hum Mol. Genet. 4:1097-1099(1995).
CC -!- FUNCTION: Could act as a modulator of transcription.
CC Expressed at low level in thymoid. Not detected in tonsil, colon,
CC lung, spleen, prostate, kidney, striated and smooth muscles.
CC Detected in rhabdomyosarcoma and fibrosarcoma cell lines. Not
CC detected in mesenchymal and epithelial cell lines.
CC -!- DISEASE: Participates in a t(X;18)(p11.2;q11.2) chromosomal
CC translocation that is specifically found in more than 80% of
CC synovial sarcoma and produces the SSX1-SSX1 or SSX1-SSX2 fusion
CC products. These hybrid proteins are probably responsible for
CC transforming activity. Heterogeneity in the position of the
CC breakpoint can occur (low frequency).
CC -!- SIMILARITY: Belongs to the SSX family.
CC -!- SIMILARITY: Contains 1 KRAB-related domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X86174; CAA60110.1; -.
CC EMBL; BC001003; AAH01003.1; -.
DR

```

```
DR EMBL; S79325; AAB5378.1; -.
DR PIR; S55057; S55057.
DR Genew; HGNC:11335; SSX1.
DR MIM; 312820; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003714; F:transcription co-repressor activity; TAS.
DR InterPro; IPR001909; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS50806; KRAB RELATED; 1.
DR Chromosomal translocation; Proto-oncogene; Multigene family;
KW Transcription regulation.
FT DOMAIN 20 83 KRAB-RELATED.
FT SITE 62 63 BREAKPOINT FOR TRANSLATION TO FORM THE
FT SITE 110 111 SSXT-SSX1 FUSION PROTEIN (RARE).
FT SITE 110 111 BREAKPOINT FOR TRANSLATION TO FORM THE
FT SITE 110 111 SSXT-SSX1 FUSION PROTEIN.
SQ SEQUENCE 188 AA; 21931 MW; E440D1B2AE3A9F7 CRC64;

Alignment Scores:
Pred. No.: 8,966-70 Length: 188
Score: 784.00 Matches: 149
Percent Similarity: 84.04% Conservative: 9
Best Local Similarity: 79.26% Mismatches: 30
Query Match: 76.43% Indels: 0
DB: 1 Gaps: 0

US-09-975-856-1 (1-576) x SSX1_HUMAN (1-188)
QY 1 ATGAACGAGAGCGCGCTTTTCAGAGAGACCCAGGATGATGCTCAATATCAGAGAG 60
Db 1 MetlanGlyAspAspThrPheAlaLysArgProArgAspAlaLysAlaSerGluLys 20
QY 61 TTACGAAAGCGCTTCGATGATATTCCTCAAGAAAGAGTGGGAAAGATG 120
Db 21 ArgSerLysAlaPheAspAlaThrTyrPheSerLysLysGluTyrLysMet 40
QY 121 AAATCTCGGAGAAATCGTCTATGTATGATGAGCTAACTATGAGGTGATGACTAAA 180
Db 41 LysTyrSerGluLysLysSerTyrValTyrMetLysArgAsnTyrLysAlaMetThrLys 60
QY 181 CTAGTTTCAAGGTCACCTCCACCTTTCATGCGTACTGAGTCAAGGTCGAGCTCCACT 240
Db 61 LeuGlyPheLysValThrLeuProPheMetCysAsnLysGlnAlaThrAspPheGln 80
QY 241 GGGATGATTTTGGTACGATCGCAACACACAGGAATCAGGTGACGTCCTCAGATGACT 300
Db 81 GlyAsnAspPheAspAsnAspHisAsnArgArgGlnValGluHisProGlnMetThr 100
QY 301 TTCGCGCAGCTCCAGAGATCTTCCGAGATCATGCCAGAGAGCCAGCAGAGAGAA 360
Db 101 PheGlyArgLeuHisArgGileProlLysLysMetProLysLysProAlaGluAspGlu 120
QY 361 AATGTTTGAAGAGTCCAGAGCATCTGCGCCACAAATGATGGAAACAGCTGTC 420
Db 121 AsnAspSerLysGlyValSerGluAlaSerGlyProGlnAsnAspGlyLysGlnLeuHis 140
QY 421 CCCCGGGAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGAGGGG 480
Db 141 ProProGlyLysAlaAsnLysSerGluLysLysAsnLysArgSerGlyProLysArgGly 160
QY 481 AAACATGCTCGACCCAGACTCGGTGAGAGAAAGCAGCTGGTGGTTTATGAGAGATC 540
Db 161 LysHisAlaThrThrHisArgLeuArgGluArgLysGlnLeuValIleTyrGluGluLe 180
QY 541 AGCGACCTCGAGGAGATGAGAG 564
Db 181 SerAspProGluGluAspAspGlu 188

RESULT 5
SSX5_HUMAN
ID -SSX5_HUMAN STANDARD; PRT; 188 AA.
```

```
AC O60225; Q96AW3;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SSX5 protein.
GN SSX5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98021352; PubMed=9378559;
RA Gure A.O., Tuereci O., Sahin U., Tsang S., Scanlan M.J., Jager E.,
RA Knuth A., Pfeundschnig M., Old L.J., Chen Y.-T.;
RT "SSX: a multigene family with several members transcribed in normal
RT testis and human cancer";
RL Int. J. Cancer 72:965-971(1997).
RN [2]
SEQUENCE FROM N.A. (ISOFORM 2).
RP TISSUE=Skin;
RX MEDLINE=22380257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Could act as a modulator of transcription.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O60225-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O60225-2; Sequence=VSP_006274;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the SSX family.
CC -!- SIMILARITY: Contains 1 KRAB-related domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U00842; AAC05821.1; -.
CC EMBL; BC016640; AAH16640.1; -.
CC Genew; HGNC:11339; SSX5.
CC MIM; 300327; -.
CC InterPro; IPR001909; KRAB.
CC InterPro; IPR003655; KRAB-related.
CC Pfam; PF01352; KRAB; 1.
CC SMART; SM00349; KRAB; 1.
CC PROSITE; PS50806; KRAB RELATED; 1.
CC Multigene family; Transcription regulation; Alternative splicing.
FT DOMAIN 20 83 KRAB-RELATED.
FT VARSPPLIC 23 23 K -> KHPMQVCORGIHLVNLSPFWKVGREPASSIKALIC
```



```

Db      36  LysThrLeuTyrAsnAspValMetGlnAspIleTyrGluThrValIleSerLeuGlyLeu 55
QY      190  AAGGTACACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCACGGGAATGAT 249
Db      56  LysLeuLys-----AsnAsp 60
QY      250  TTGTGTAACGATCGA-----264
Db      61  ThrGlyAsnAspHisProIleSerValSerThrSerGluIleGlnThrSerGlyCysGlu 80
QY      265  ---AACCACAGAAATCAGGTTCAGCTCCTCAGATGACTTTCGGC-----306
Db      81  ValSerLysLysThrArgMetLysIleAlaGlnLysThrMetGlyArgGluAsnProGly 100
QY      307  -----AGCTCCAG-----AGAACTTCCCGAAGATCATGCCCAAGAGCCA 348
Db      101  AspThrHisSerValGlnLysTrpHisArgAlaPheProArgLysLysArgLysLysPro 120
QY      349  GCAGAGGAGAAATGTTTGAGGAGTCCAGAGGATCTGGCCCAAAATGATGGG 408
Db      121  AlaThrCysLysGlnGluLeuProLysLeuMetAspLeuHisGlyLysGlyProThrGly 140
QY      409  AAACAG 414
Db      141  GluLys 142

RESULT 7
ID      2398 HUMAN STANDARD; PRT; 642 AA.
AC      Q8TD17; Q8TD18; Q9P2K7; Q9UDV8;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DE      Zinc finger protein 398 (Zinc finger DNA binding protein p52/p71).
GN      ZNF398 OR ZER6 OR KIAA1339.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI TaxID:9606;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX      MEDLINE=21883946; PubMed=11779858;
RA      Conroy A.T., Sharma M., Holz A.E., Wu C., Sun Z., Weigel R.J.;
RT      "A novel zinc finger transcription factor with two isoforms that are
RL      differentially repressed by estrogen receptor-alpha.";
RL      J. Biol. Chem. 277:9326-9334(2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Bemis G., Langston Y., Tucci S.;
RL      Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RX      MEDLINE=22388257; PubMed=12477932;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zienberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
```

```

RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [4]
RP      SEQUENCE OF 234-642 FROM N.A.
RX      MEDLINE=20181126; PubMed=10718198;
RA      Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
RT      "Prediction of the coding sequences of unidentified human genes. XVI.
RT      The complete sequences of 150 new cDNA clones from brain which code
RL      for large proteins in vitro.";
RL      DNA Res. 7:65-73(2000).
CC      !- FUNCTION: Function as a transcriptional activator.
CC      !- SUBCELLULAR LOCATION: Nuclear (Probable).
CC      !- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=2;
CC      Name=1; Synonyms=p71;
CC      IsoId=Q8TD17-1; Sequence=Displayed;
CC      Name=2; Synonyms=p52;
CC      IsoId=Q8TD17-2; Sequence=VSP 006926;
CC      !- INDUCTION: By estrogen receptor alpha.
CC      !- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC      FINGER PROTEINS.
CC      !- SIMILARITY: Contains 1 KRAB domain.
CC      !- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC      gene model prediction.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AY049744; AAK32789.1; -
DR      EMBL; AY049743; AAK32788.1; -
DR      EMBL; AY004890; AAD45824.1; ALT_SEQ.
DR      EMBL; BC043295; AAH43295.1; -
DR      EMBL; AB037760; BAA92577.1; -
DR      TRANSFAC; T05129; -
DR      Genew; HGNC:18373; ZNF398.
DR      GO; GO:0005634; C:nucleus; NAS.
DR      GO; GO:0016563; F:transcriptional activator activity; NAS.
DR      GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR      InterPro; IPR001309; KRAB.
DR      InterPro; IPR007087; Znf_C2H2.
DR      Pfam; PF01352; KRAB; 1.
DR      Pfam; PF00096; zf-C2H2; 8.
DR      ProDom; PD000003; Znf_C2H2; 2.
DR      SMART; SM00349; KRAB; 1.
DR      SMART; SM00355; Znf_C2H2; 8.
DR      PROSITE; PS00805; KRAB; 1.
DR      PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR      PROSITE; PS0157; ZINC_FINGER_C2H2_2; 8.
KW      Transcription regulation; Activator; DNA-binding; Zinc-finger;
KW      Metal-binding; Nuclear protein; Repeat; Alternative splicing.
FT      DOMAIN 143 214
FT      ZN_FING 343 364 C2H2-TYPE (ATYPICAL).
FT      ZN_FING 392 392 C2H2-TYPE (DEGENERATE).
FT      ZN_FING 398 420 C2H2-TYPE.
FT      ZN_FING 427 449 C2H2-TYPE.
FT      ZN_FING 455 477 C2H2-TYPE.
FT      ZN_FING 483 505 C2H2-TYPE.
FT      ZN_FING 511 533 C2H2-TYPE.
FT      ZN_FING 539 561 C2H2-TYPE.
FT      ZN_FING 567 590 C2H2-TYPE.
FT      VARSPLIC 1 171 Missing (in isoform 2).
FT      /FTId=VSP 006926.
SQ      SEQUENCE 642 AA; 71311 MW; 69AA38FCDB4FP633 CRC64;
Alignment Scores:
Pred. No.: 0.0593 Length: 642
Score: 101.00 Matches: 31
Percent Similarity: 41.60% Conservative: 21
Best Local Similarity: 24.80% Mismatches: 41
```



```

Query Match: 9.85% Indels: 32
DB: 1 Gaps: 5
US-09-975-856-1 (1-576) x Z317_HUMAN (1-642)

QY 58 AAGTTACGAAGGCTTCGATGATATGCGCAATACCTCTTAAGAAAGAGTGGAAAAG 117
D 140 LysValProValAlaPheAspValSerThrPheSerThrProGluTrpGluLys 159
QY 118 ATGAATCTCCGGAGAAATCGTCTATGTG---TATATGAAGCTAAACTATGAGTTCATG 174
D 160 LeuGluGluTrpGlnLysGluLeuTyLysAsnIleMetLysGlyAsnTyLysSerLeu 179
QY 175 ACTAACTAGTTCGAAGTCAACCTCCACCTTCATCGTGTAGTAAACGGGCTGCAGAC 234
D 180 IleSerMetAspTyLysAlaIleAsnGlnProAspValLeuSerGlnIleGlnProGluGly 199
QY 235 TTCACGGGATGATTTTGGTAAAGTCAACACACAGGAAATCAGGTTGACGCTCCTCAG 294
D 200 GluHisAsnThrGlu-----AspGlnAlaGlyProGluGluSerGlu----- 213
QY 295 ATGACTTTCCGCGACCTCCAGAGAATCTCCGGAAGATCATGCCCAAGAGCCAGCAGAG 354
D 214 ----- 220
QY 355 GAAGAAATCGTTG----- 381
D 221 GluPro---GlyIleSerThrSerAspIleLeuSerTrpIleLysGlnGluGluPro 239
QY 382 GAGGATCTGGCCCA 396
D 240 GlnValGlyAlaPro 244

RESULT 8
Z317_HUMAN STANDARD; PRT; 595 AA.
AC Q96P06; Q96P01; Q96P02;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 317.
GN ZNF317.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
[1]
RN RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4).
RX MEDLINE=21547764; PubMed=11688974;
RA Takaehima H., Nishio H., Wakao H., Nishio M., Koizumi K., Oda A.,
RA Koike T., Sawada K. K.;
RT "Molecular cloning and characterization of a KRAB-containing zinc
RT finger protein, ZNF317, and its isoforms."
RL Biochem. Biophys. Res. Commun. 288:771-779(2001).
CC -!- FUNCTION: May function as a transcription factor. May play an
CC important role in erythroid maturation and lymphoid proliferation.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=2; Synonyms=ZNF317-2;
CC IsoId=Q96P06-1; Sequence=Displayed;
CC Name=1; Synonyms=ZNF317-1;
CC IsoId=Q96P06-2; Sequence=VSP_006916;
CC Name=3; Synonyms=ZNF317-3;
CC IsoId=Q96P06-3; Sequence=VSP_006915, VSP_006916;
CC Name=4; Synonyms=ZNF317-4;
CC IsoId=Q96P06-4; Sequence=VSP_006915;
CC -!- TISSUE SPECIFICITY: Isoforms 1 and 3 are ubiquitously expressed.
CC Isoforms 3 and 4 are expressed only in lymphocytes, spleen, and
CC lung.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 1 KRAB domain.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; AF275255; AAL29188.1; -
DR EMBL; AF148135; AAL29182.1; -
DR EMBL; AF307096; AAL29190.1; -
DR EMBL; AF307097; AAL29191.1; -
DR PIR; JC7779; JC7779.
DR Genew; HGNC:13507; ZNF317.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR001909; Znf_C2H2.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 13.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 13.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 13.
KW Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein;
KW DNA-binding; Repeat; Alternative splicing.
FT DOMAIN 57 128 KRAB.
FT ZN FING 222 244 C2H2-TYPE.
FT ZN FING 250 272 C2H2-TYPE.
FT ZN FING 278 300 C2H2-TYPE.
FT ZN FING 306 328 C2H2-TYPE.
FT ZN FING 334 356 C2H2-TYPE.
FT ZN FING 362 384 C2H2-TYPE.
FT ZN FING 390 412 C2H2-TYPE.
FT ZN FING 418 440 C2H2-TYPE.
FT ZN FING 446 468 C2H2-TYPE.
FT ZN FING 474 496 C2H2-TYPE.
FT ZN FING 502 524 C2H2-TYPE.
FT ZN FING 530 552 C2H2-TYPE.
FT ZN FING 558 580 C2H2-TYPE.
FT VARSPLIC 1 85 Missing (in isoform 3 and isoform 4).
FT VARSPLIC 97 128 /FTId=VSP_006915.
FT VARSPLIC 97 128 Missing (in isoform 1 and isoform 3).
FT /FTId=VSP_006916.
SQ SEQUENCE 595 AA; 67941 MW; 9AFF5742A34EB959 CRC64;

Alignment Scores:
Pred. No.: 0.289 Length: 595
Score: 94.00 Matches: 32
Percent Similarity: 47.62% Conservative: 18
Best Local Similarity: 30.48% Mismatches: 51
Query Match: 9.17% Indels: 4
DB: 1 Gaps: 3

US-09-975-856-1 (1-576) x Z317_HUMAN (1-595)
QY 73 TTCGATGATATGCGCAATACCTCTTAAGAAAGTGGAAAAGATGAAATCCTCGGAG 132
D 59 PheGlnAspValAlaValAspPheThrGluLysGluTrpProLeuLeuAspSerSerGln 78
QY 133 AAAATCGTCTATGCTATATGAAGCTA---AACTATGAGGTGATGACTAACTAGGTTTC 189
D 79 ArgLysLeuTyLysAspValMetLeuGluAsnTyLysSerAsnLeuThrSerLeuGlyTyr 98
QY 190 AAGGTCACCTCCACCTTTTCATG---CGTAGTAAACGGGCTGCAGACTTCCACGGGAAT 246
D 99 GlnValGlyLysProSerLeuLeuSerHisLeuGluGlnGluGluProArgThrGlu 118
QY 247 GATTTGGTAAAGTCAACACACAGGAAATCAGGTTGACGCTCCTCAGATGACTTTCGGC 306
D 119 GluArgGlyAlaHisGlnGlyAlaCysAlaAspTrpGluThrProSerLysThrLysTrp 138
QY 307 AGCCTC-----CAGAGATCTTCCCGAAGATCATGCCCAAGAGCCAGCAGAGAGAA 360

```

```

Db 139 SerLeuLeuMetGluAspIlePheGlyIysGluThrProSerGlyValThrMetGluArg 158
QY 361 AATGGTTTGAAGGAA 375
Db 159 AlaGlyLeuGlyGlu 163

RESULT 9
PK3G_RAT
ID PK3G_RAT STANDARD; PRT; 1505 AA.
AC O70173;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Phosphatidylinositol-4-phosphate 3-kinase C2 domain-containing gamma
DE polypeptide (EC 2.7.1.154) (Phosphoinositide 3-Kinase-C2-gamma)
DE (PTdIns-3-kinase C2 gamma) (PI3K-C2gamma).
GN PI3K2G.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
RP STAGE.
RC TISSUE=Regenerating liver;
RX MEDLINE=98184888; PubMed=9516481;
RA Ono F., Nakagawa T., Saito S., Owada Y., Sakagami H., Goto K.,
RA Suzuki M., Matsuno S., Kondo H.;
RT "A novel class II phosphoinositide 3-kinase predominantly expressed in
RT the liver and its enhanced expression during liver regeneration.";
RL J. Biol. Chem. 273:7731-7736 (1998).
CC -!- FUNCTION: IN VITRO, PHOSPHORYLATES PTDINS AND PTDINS4P BUT NOT
CC PTDINS(4,5)P2.
CC -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 4-
CC phosphate = ADP + 1-phosphatidyl-1D-myo-inositol 3,4-bisphosphate.
CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN NORMAL LIVER. HIGH
CC LEVELS ALSO FOUND IN REGENERATING LIVER. VERY LOW LEVELS FOUND IN
CC HEART AND TESTIS.
CC -!- DEVELOPMENTAL STAGE: HIGHER LEVELS OF EXPRESSION FOUND IN ADULT
CC LIVER THAN IN FETAL LIVER.
CC -!- SIMILARITY: Belongs to the PI3/P14-kinase family.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 1 pbox homology (PX) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB009636; BAA25634.1; -.
CC InterPro; IPR008938; ARM.
CC InterPro; IPR000008; C2.
CC InterPro; IPR008973; C2_CaLB.
CC InterPro; IPR000403; PI3_P14_kinase.
CC InterPro; IPR002420; PI3K_C2_kinase.
CC InterPro; IPR000341; PI3K_ras_bind.
CC InterPro; IPR001263; PI3Ka.
CC InterPro; IPR001693; PX.
CC Pfam; PF00168; C2; 1.
CC Pfam; PF00454; PI3_P14_Kinase; 1.
CC Pfam; PF00792; PI3K_C2; 1.
CC Pfam; PF00794; PI3K_rbd; 1.
CC Pfam; PF00613; PI3Ka; 1.
CC Pfam; PF00787; PX; 1.
CC SMART; SM00239; C2; 2.
CC SMART; SM00142; PI3K_C2; 1.
CC SMART; SM00144; PI3K_rbd; 1.
CC SMART; SM00145; PI3Ka; 1.

```

```

DR SMART; SM00146; PI3Kc; 1.
DR SMART; SM00312; PX; 1.
DR PROSITE; PS00499; C2_DOMAIN_1; FALSE_NEG.
DR PROSITE; PS00004; C2_DOMAIN_2; FALSE_NEG.
DR PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR PROSITE; PS02290; PI3_4_KINASE_3; 1.
DR PROSITE; PS01195; PX; 1.
KW Transferase; Kinase; Membrane; Multigene family.
FT DOMAIN 976 1240 PI3K/PI4K.
FT DOMAIN 1259 1371 PX.
FT DOMAIN 1402 1499 C2_DOMAIN.
SQ SEQUENCE 1505 AA; 170974 MW; 5ED4C2239968C4B2 CRC64;

Alignment Scores:
Pred. No.: 0.95 Length: 1505
Score: 89.50 Matches: 33
Percent Similarity: 42.24% Conservative: 16
Best local Similarity: 28.45% Mismatches: 50
Query Match: 8.89% Indels: 17
DB: 1 Gaps: 5

US-09-975-856-1 (1-576) x PK3G_RAT (1-1505)
QY 490 AGGCATGTTTCCCTTTGGGTCCAGATGCTCTTTAACTCTTCCAGAGTACTTGAT 431
Db 83 ArgHisPheAsnGluPheThrSerGln-----SerProHisPheSerGln 97
QY 430 TTCGGGGGGGCACAGCGTTTCCCATCATTTTGTGGGCCAGATCGCTCTGSCATCTCT 371
Db 98 LeuProPheGlyLysAlaSerAla-----IleGlyPheAsnProAlaValLeuPro 114
QY 370 TCAACACCATTTTCTCTCTCTGCTGCTTCTTGGGCATGATCTTCGGGAAGATCTCTCGA 311
Db 115 AlaHisGlnPheIleHisGluGlyAlaSerTrpArgAsnProThrArgLysTyrHisGly 134
QY 310 GCCTGCCGAAAGTCATCTGAGGACGCTTCAACTGATCTCTGTGGTTTGCATCGT---TAC 254
Db 135 GlyGluAspProArgPheSerAlaLeuThrProSerSerThrGlyLeuAspLysCysHis 154
QY 253 CAAATCATCTCCGTGGGAAGCTGCAGCCGCTTACTAGCATGAAAGTGGAGGTGA 194
Db 155 GlnGlnGlnSerGlyThrGluHisCysAsnTyrTyrVal-----Glu 169
QY 193 CTTTGAACACCTAGTGTAGTACCTCATGACCTCATAGTTTGTAGCTTCATATACA 146
Db 170 ProGluAsnAsnVal-----ProHisTyrSerProTyrSer 182

RESULT 10
FM01_CANFA
ID FM01_CANFA STANDARD; PRT; 531 AA.
AC Q951A2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dimethylalanine monooxygenase [N-oxide forming] 1 (EC 1.14.13.8)
DE (Hepatic flavin-containing monooxygenase 1) (FM0 1) (Dimethylalanine
DE oxidase 1).
GN FM01.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Lattard V., Longin-Sauvageon C., Lachuer J., Buronfosse T., Benoit E.;
RT "Cloning, sequencing and tissue dependent expression of FM01 and FM03
RT in the dog.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: This protein is involved in the oxidative metabolism of
CC a variety of xenobiotics such as drugs and pesticides.
CC -!- CATALYTIC ACTIVITY: N,N-dimethylaniline + NADPH + O(2) = N,N-

```

```

CC dimethylamine N-oxide + NADP(+) + H(2)O.
CC -I- COFACTOR: FAD (By similarity).
CC -I- SUBCELLULAR LOCATION: Mitochondrion.
CC -I- TISSUE SPECIFICITY: Liver.
CC -I- SIMILARITY: Belongs to the FMO family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR ENBL; AF384053; AAK97433.1;
DR InterPro: IPR000759; Admrdx_reductase.
DR InterPro: IPR001327; FAD_Pyridoxal.
DR InterPro: IPR000960; FAD_Pyridoxal.
DR Pfam: PF00743; FMO-like; 1.
DR PRINTS: PR00419; ADXRUTASE.
DR PRINTS: PR00368; FADPR.
DR PRINTS: PR00370; FMOXYGENASE.
KW Oxidoreductase; Monooxygenase; NADP; Flavoprotein; FAD; Microsome;
KW Transmembrane; Multigene family; Acetylation.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 8 13 ACETYLATION (BY SIMILARITY).
FT NP_BIND 190 195 FAD (ADP PART) (POTENTIAL).
FT NP_BIND 190 195 NADP (POTENTIAL).
SQ SEQUENCE 531 AA; 59927 MW; 9F3458484540521F CRC64;

Alignment Scores:
Pred. No.: 0.996 Length: 531
Score: 88.50 Matches: 33
Percent Similarity: 38.21% Conservative: 14
Best Local Similarity: 26.83% Mismatches: 36
Query Match: 8.79% Indels: 40
DB: 1 Gaps: 4

US-09-975-856-1 (1-576) x FMO1_CANFA (1-531)
QY 565 ACTGTCATCTCTCCAGGTCCTGATCTCTTCATTAACACACAGCTGCTTCTCTCAC 506
Db 43 ThrGluHisValGluGluGlyArgAlaSerLeuTyrLysSerVal----- 57
QY 505 GCAGTCGTGGGTCGAGGATGTTCCCTTTGGGTCCAGATGCTGTTAACTCTT 446
Db 58 -----ValSerAsnSerCysLysGluMetSerCys----- 67
QY 445 CCAAGGTACTTGATTTCCCGGGGGGCACAGCTGTTTCCCATCATTTTGGGCCAGATG 386
Db 68 -----TyrSerAspPhePro-----PheProGluAspTyr 77
QY 385 CCTCTGGCACTTCTTCAACCACTTTCTTCCTCTGCTGCTTCTTGGGCAATCTTCG 326
Db 78 ProAsnTyrValProAsnSerGlnPheLeuGluTyrLeuLysMetTyrAla-Asn----- 95
QY 325 GGAAGATCTCTGGAGGTCGCGAAGATCATCTGAGGAGCTTCAACTGATCTCTGCT 266
Db 96 -----ArgPheSerLeuLeuLysCys11 103
QY 265 TCGATCGTTTACCAAAATCATCTCCGTTGGAAGTCTGACCGCTTTTACTACGCATGAAG 206
Db 103 eArgPheLysThrLysValCysLysValThrLysCysProAspPheThrValThrGlyG 123
QY 205 GTGGAG 199
Db 123 nTyrGlu 125

RESULT 11
REPL_MOUSE
ID REPL_MOUSE STANDARD; PRT; 743 AA.
AC O54916; Q8C9J9; Q99LR8;
DT 10-OCT-2003 (Rel. 42, Created)

```

```

DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE RalBP1 associated Eps domain containing protein 1 (RalBP1-interacting
DE protein 1).
DE REPS1.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RC TISSUE=Muscle;
RX MEDLINE=98058900; PubMed=9395447;
RA Yamaguchi A., Urano T., Goi T., Feig L.A.;
RT "An eps homology (EH) domain protein that binds to the ral-GTPase
RL J. Biol. Chem. 272:31230-31234 (1997).
RN [2]
RP SEQUENCE OF 8-743 FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN [3]
RP SEQUENCE OF 447-743 FROM N.A. (ISOFORM 1).
RC TISSUE=Breast tumor;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."

```

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[4]
 RN STRUCTURE BY NMR OF 227-318.
 RX MEDLINE=12185759; PubMed=11389591;
 RA Kim S., Cullis D.N., Feig L.A., Balleja J.D.;
 RT "Solution structure of the Rep1 EH domain and characterization of
 RT its binding to NPF target sequences."
 RL Biochemistry 40:6776-6785(2001).
 CC !- FUNCTION: May coordinate the cellular actions of activated EGF
 CC receptors and Ral-GTPases.
 CC !- SUBUNIT: Homodimer (Potential). Interacts with RALBP1, CRK and
 CC GRB2. Binding to RALBP1 does not affect its Ral-binding activity.
 CC Forms a complex with the SH3 domains of CRK and GRB2 which may
 CC link it to an EGF-responsive tyrosine kinase.
 CC !- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=O54916-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O54916-2; Sequence=VSP_007956, VSP_007957;
 CC Notes=Due to intron retention. No experimental confirmation
 CC available;
 CC !- TISSUE SPECIFICITY: Expressed in all tissues examined. The highest
 CC level expression was found in the kidney and testis.
 CC !- PTM: EGF stimulates phosphorylation on tyrosine residues.
 CC !- SIMILARITY: Contains 1 EH domain.
 CC !- SIMILARITY: Contains 1 EF-hand calcium-binding domain.
 CC !- CAUTION: Ref.3 sequence differs from that shown due to a
 CC frameshift in position 719.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF031939; AAB94736.1; -;
 DR EMBL; AK041967; BAC31117.1; ALT INIT.
 DR EMBL; BC022256; AAB02256.1; ALT_FRAME.
 DR PIR; T09173; T09173.
 DR PDB; 1RT6; 18-JUL-01.
 DR MGD; MGI:1196373; Reps1.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR000261; EPS15_homology.
 DR Pfam; PF00036; efhand; 1.
 DR SMART; SM00027; EH; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00031; EH; 1.
 KW Calcium-binding; Coiled coil; Phosphorylation; Alternative splicing;
 KW 3D-structure.
 FT DOMAIN 233 324 EH.
 FT CA_BIND 279 290 EF-HAND (POTENTIAL).
 FT DOMAIN 488 551 PRO-RICH.
 FT DOMAIN 599 743 INTERACTION WITH RALBP1.
 FT DOMAIN 692 738 COILED COIL (POTENTIAL).
 FT MOD_RES 236 236 PHOSPHORYLATION (POTENTIAL).
 FT VARSPPLIC 368 402 OWETFSRSSSQTLTQPSNIAPADPTAIYHPV ->
 FT VSKTSLSLLEISLFTGRSKQDRFTAGVLYAHTP (in
 FT isoform 2).
 FT /FTId=VSP_007956.
 FT Missing (in isoform 2).
 FT /FTId=VSP_007957.
 SQ SEQUENCE 743 AA; 80598 MW; 25510D1I254CF4A6 CRC64;
 Alignment Scores:
 Pred. No.: 1.18 Length: 743
 Score: 88.00 Matches: 40
 Percent Similarity: 38.92% Conservative: 25
 Best Local Similarity: 23.95% Mismatches: 41
 Query Match: 8.59% Gaps: 61
 DB: 1 Indels: 9

US-09-975-856-1 (1-576) x REPI_MOUSE (1-743)

QY 91 TACTTCTTAAGAAAGAGTGGGAAAAGATGAATCCTCGAGAAAATCCTATGTGTAT 150
 ||||| : : : : :
 Db 11 TyrPheGlyArgSerGlnPhe : : : : : Tyr 18
 : : : : :
 QY 151 ATGAAGCTAAACTATCAGTCACTAACTAGTTTC : : : : : 189
 : : : : :
 Db 19 IleAlaLeuLeuValAlaGlnSerGlyPheProLeuArgValGluSerIle 38
 : : : : :
 QY 190 : : : : : AAGTCACCTCCACCTTTTCATCGGTAGTAAA : : : : : 222
 : : : : :
 Db 39 AsnThrVallyAspLeuProLeuProArgPheValAlaSerLysAsnGluGlnGluSer 58
 : : : : :
 QY 223 CGGCTGCAGACTTCCCGGAATCATTTTGTACGATCGAACCCAC : : : : : 270
 : : : : :
 Db 59 ArgLeuAlaAlaSerTy-SerSerAspSerGluAsnGlnGlySerTy-SerGlyValIle 78
 : : : : :
 QY 271 : : : : : AGGAATCAGGTTGAACCTCCTCAGATGACTTTCGGCAGCTCCAG 315
 : : : : :
 Db 79 ProProProGlyArgGlyGlnVallyLysGlyProGlySerHisAspAlaValGln 98
 : : : : :
 QY 316 AGAATCTCCGAAAGATCATGCCAAGAACGCCAGCAGAGAGAAATGTTTGAAGAA 375
 : : : : :
 Db 99 : : : : : ProArgProSerAlaGluGlnGln : : : : : 106
 : : : : :
 QY 376 GTGCCAGAGGATCTGGCCCAAAATGATGGAAACAGCTGTGCCCCCGGMAATCCA 435
 : : : : :
 Db 107 : : : : : GluProAlaSerProValValSerProGlnGln : : : : : 123
 : : : : :
 QY 436 AGTACTTTGGAGAAGATTAAACAAGACATCTGACCCAAAGGGGGAACATG : : : : : 487
 : : : : :
 Db 124 HisThr-TrpArgLysHisSerArgHis : : : : : ProSerGlyGlyAsnSerGluArgPr 141
 : : : : :
 QY 488 : : : : : CCTGACCC 496 : : : : :
 : : : : :
 Db 141 oLeuThrGlyProGlyPro 147

RESULT 12
 FMO1_PIG STANDARD; PRT; 531 AA.
 AC P16549;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Dimethylaniline monooxygenase [N-oxide forming] 1 (EC 1.14.13.8)
 DE (Hepatic flavin-containing monooxygenase 1) (FMO 1) (Dimethylaniline
 DE oxidase 1).
 DE FMO1 OR FMO-1.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]_TaxID=9823;
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 136-150 AND 308-317.
 RC TISSUE=Liver;
 RX MEDLINE=90212556; PubMed=2322534;
 RA Gasser R., Tynes R.E., Lawton M.P., Korsmeyer K.K., Ziegler D.M.,
 RA Philpot R.M.;
 RT "The flavin-containing monooxygenase expressed in pig liver: primary
 RT sequence, distribution, and evidence for a single gene."
 RL Biochemistry 29:119-124 (1990).
 RN [2]
 RP SEQUENCE OF 1-13 AND 184-201, AND ACETYLATION.
 RC TISSUE=Liver;
 RX MEDLINE=90343821; PubMed=2383273;
 RA Guan S.H., Fallick A.M., Cashman J.R.;
 RT "N-terminus determination: FAD and NADP binding domain mapping of hog
 RT liver flavin-containing monooxygenase by tandem mass spectrometry."
 RL Biochem. Biophys. Res. Commun. 170:937-943 (1990).
 RN [3]
 RP SEQUENCE OF 185-207.

ENBL; X07290; CAA30269.1; --	
PIR; S00754; S00754.	
Genew; HGNC:13104; ZNF38.	
Genew; HGNC:13089; ZNF3.	
MIM; 601261; --	
MIM; 194510; --	
GO; GO:0005634; C:nucleus; IC.	
GO; GO:0003700; F:transcription factor activity; NAS.	
GO; GO:0008270; F:zinc ion binding; NAS.	
GO; GO:0045321; P:cell activation; NAS.	
GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.	
InterPro; IPR001909; KRAB.	
InterPro; IPR007087; Znf_C2H2.	
InterPro; IPR007086; Znf_C2H2_sub.	
Pfam; PF01352; KRAB; 1.	
Pfam; PF00096; zf-C2H2; 8.	
PRINTS; PR00048; ZINC_FINGER.	
ProDom; PD000003; Znf_C2H2; 8.	
SMART; SM00349; KRAB; 1.	
SMART; SM00355; Znf_C2H2; 8.	
PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.	
PROSITE; PS0157; ZINC_FINGER_C2H2_2; 8.	
PROSITE; PS0805; KRAB; 1.	
Transcription regulation; Activator; Zinc-finger; DNA-binding;	
Repeat; Metal-binding; Nuclear protein; Differentiation.	
DOMAIN 51 123	KRAB.
ZN_FING 200 227	C2H2-TYPE.
ZN_FING 228 255	C2H2-TYPE.
ZN_FING 256 283	C2H2-TYPE.
ZN_FING 284 311	C2H2-TYPE.
ZN_FING 312 339	C2H2-TYPE.
ZN_FING 340 367	C2H2-TYPE.
ZN_FING 368 395	C2H2-TYPE.
ZN_FING 396 423	C2H2-TYPE.
ZN_FING 421 421	L -> P (IN REF. 2).
CONFLICT 252 256	GKPY -> IRDSG (IN REF. 5).
CONFLICT 336 387	GKPYECNCKAKSQSSHLVQHORIHTGKPYECMECCGK
CONFLICT 336 387	FTYSGLLQHQ -> EALPTVTLIRLLPSPDPTVINEAAF
CONFLICT 336 387	PAESIATLIALIRLFCVHSLMPKKV (IN REF. 3).
SEQUENCE 446 AA; 50932 MW; 67A626807304782 CRC64;	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 31, 2004, 13:40:37 ; Search time 47 Seconds
(without alignments)
7733.553 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 1025

Sequence: 1 ATGACGGAGACACGCCCTT.....ATGACGAGTACTCCCTCG 576

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q=/cg2_1/USPTO.spool_p/US0975856/runat_31032004_132903_14748/app_query.fasta_1.775
-DB=SPTRMBL 25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=opt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0975856 -CGN_1_1_86 -runat_31032004_132903_14748 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTRMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	685.5	66.9	223	4	Q9BU88 homo sapien

ALIGNMENTS

RESULT 1	ID	Q9BU88	PRELIMINARY;	PRT;	223 AA.
Q9BU88	AC	Q9BU88;			
DT	01-JUN-2001	(TRENBLrel. 17, Created)			
DT	01-JUN-2001	(TRENBLrel. 17, Last sequence update)			
DT	01-OCT-2003	(TRENBLrel. 25, Last annotation update)			
DE		Similar to synovial sarcoma, X breakpoint 2.			
OS		Homo sapiens (Human).			
OC		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX		NCBI_TaxID=9606;			
RN	[1]	SEQUENCE FROM N.A.			
RP		TISSUE=Placenta;			
RA		Strausberg R.;			
RL		Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL;	BC002818; AA02818.1; -			
DR	GO;	GO:0005622; C:intracellular; IEA.			
DR	GO;	GO:0003676; F:nucleic acid binding; IEA.			
DR	GO;	GO:0006355; P:regulation of transcription, DNA-dependent; IEA.			
DR	InterPro;	IPR001909; KRAB.			
DR	InterPro;	IPR003655; KRAB_related.			
DR	Pfam;	PF01352; KRAB; 1			
DR	SMART;	SM00349; KRAB; 1.			

Q9brw7 homo sapien
Q96q11 homo sapien
Q96q10 homo sapien
Q9y444 homo sapien
Q9wz29 homo sapien
Q9nzka4 homo sapien
Q8c5z3 mus musculus
Q8c5z3 mus musculus
Q75101 homo sapien
Q81zh0 homo sapien
Q81zg9 homo sapien
Q81zg8 homo sapien
Q81zg7 homo sapien
Q81zg6 homo sapien
Q81zg5 homo sapien
Q81zh1 homo sapien
Q86td5 homo sapien
Q60290 homo sapien
Q96mn0 homo sapien
Q8n2j5 homo sapien
Q9uld5 homo sapien
Q8bfs8 mus musculus
Q9hr87 botrytis ci
Q8qzz2 mus musculus
Q8n393 homo sapien
Q9wv10 mus musculus
Q81w91 homo sapien
Q8c393 mus musculus
Q8bv16 mus musculus
Q8c887 mus musculus
Q8bpj3 mus musculus
Q80tc5 mus musculus
Q8bvho mus musculus
Q9p215 homo sapien
Q9udv5 homo sapien
Q8c964 methanosarc
Q8tsp6 methanosarc
Q8tuh2 methanosarc
Q8tyt7 methanosarc
Q8tna0 methanosarc
Q8cmcs methanosarc
Q8tl32 methanosarc
Q8til4 methanosarc

```
DR PROSITE; P50806; KRAB RELATED; 1.
SQ SEQUENCE 223 AA; 25173 MW; 2BF9E1FFA4D58094 CRC64;

Alignment Scores:
Pred. No.: 1,15e-62 Length: 223
Score: 685.50 Matches: 137
Percent Similarity: 66.07% Conservative: 11
Best Local Similarity: 61.16% Mismatches: 26
Query Match: 66.88% Indels: 50
DB: 4 Gaps: 1

US-09-975-856-1 (1-576) x Q9BU88 (1-223)
QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGAGACCCAGGAGATGCTCTCAATATATCAGAGAAG 60
Db 1 MetAsnGlyAspAlaPheAlaArgProThrValGlyAlaGlnLeuProGluLys 20
QY 61 TTACGAAAGCCCTCGATGATATGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
Db 21 IleGlnLysAlaPheAspAlaLysTyrPheSerLysGluLysMet 40
QY 121 AATCTCTCGAGAAATCGTCTATGTATGATCAAGCTAACTATCAGGTCTGACTATAA 180
Db 41 LysAlaSerGluLysLeuPheTyrValTyrMetLysArgLysTyrGluAlaMetThrLys 60
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGTAGTAACCGGCTGCAGACTTCCAC 240
Db 61 LeuGlyPheLysAlaThrLeuProPheMetCysAsnLysArgAlaGluAspPheGln 80
QY 241 GGAATGATTTTGGTACGATGCAACACACAGAAATCAGTTGAAAGTCTCGAGACT 300
Db 81 GlyAsnAspLeuAspAsnAspProAsnArgGlyAsnGlnValGluArgProGlnMetThr 100
QY 301 TTCCGCGAGCTCCAGAGAAATCTTCCGAGATCATGCCAAGAGCCAGAGGAAGAA 360
Db 101 PheGlyArgLeuGlnGlyIleSerProLysIleMetProLysLysProAlaGluGly 120
QY 361 AATGTTTGAAGAGTGCAGAGGATCTGGCCCAAAATATGATGGAAACAGCTGTGC 420
Db 121 AsnAspSerGluGluValProGluAlaSerGlyProGlnAsnAspGlyLysGluLys 140
QY 421 CCCCAGGAAATCAAGTACCTTGGAGAGATTAAACAGACATCTGA 468
Db 141 ProGlyLysProThrThrSerGluLysIleHisGluArgSerGlyAsnArgGluAla 160
QY 468 ----- 468
Db 161 GlnGluLysGluGluArgArgGlyThrAlaHisArgTrpSerSerGlnAsnThrHisAsn 180
QY 468 ----- 468
Db 181 IleGlyArgPheSerLeuSerThrSerMetGlyAlaValHisGlyThrProLysThrIle 200
QY 469 -----CCCAAAAGGGGAAACATGCTGGACCCACACAGACTGCTGAGAGA 513
Db 201 ThrHisAsnArgAspProLysGlyGlyAsnMetProGlyProThrAspCysValArgG 220
QY 514 AAGCAGCTGG 523
Db 220 uAsnSerTip 223

RESULT 2
Q9BRW7 PRELIMINARY; PRT; 170 AA.
AC Q9BRW7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Synovial sarcoma, X breakpoint 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.

[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005904; AAH05904.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; P50806; KRAB RELATED; 1.
SQ SEQUENCE 170 AA; 19457 MW; D981F807A9C7EAB1 CRC64;

Alignment Scores:
Pred. No.: 1,61e-57 Length: 170
Score: 636.00 Matches: 122
Percent Similarity: 82.05% Conservative: 6
Best Local Similarity: 78.21% Mismatches: 28
Query Match: 62.05% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q9BRW7 (1-170)
QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGATGCTCTCAATATATCAGAGAAG 60
Db 1 MetAsnGlyAspAlaPheAlaArgProThrValGlyAlaGlnLeuProGluLys 20
QY 61 TTACGAAAGCCCTCGATGATATGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
Db 21 IleGlnLysAlaPheAspAlaLysTyrPheSerLysGluLysMet 40
QY 121 AATCTCTCGAGAAATCGTCTATGTATGATCAAGCTAACTATCAGGTCTGACTATAA 180
Db 41 LysValSerGluLysIleValTyrValTyrMetLysArgLysTyrGluAlaMetThrLys 60
QY 181 CTAGTCTTCAAGGTCAACCTCCACCTTTCATCGTAGTAACCGGCTGCAGACTTCCAC 240
Db 61 LeuGlyPheLysAlaLysLeuProSerPheMetArgAsnLysArgValThrAspPheGln 80
QY 241 GGAATGATTTTGGTACGATGCAACACACAGAAATCAGTTGAAAGTCTCGAGACT 300
Db 81 GlyAsnAspPheAspAsnAspProAsnArgGlyAsnGlnValGlnArgProGlnMetThr 100
QY 301 TTCCGCGAGCTCCAGAGAAATCTTCCGAGATCATGCCAAGAGCCAGAGGAAGAA 360
Db 101 PheGlyArgLeuGlnGlyIlePheProLysIleMetProLysLysProAlaGluGly 120
QY 361 AATGTTTGAAGAGTGCAGAGGATCTGGCCCAAAATATGATGGAAACAGCTGTGC 420
Db 121 AsnValSerLysGluValProGluAlaSerGlyProGlnAsnAspGlyLysGlnLeuLys 140
QY 421 CCCCAGGAAATCAAGTACCTTGGAGAGATTAAACAGACATCTGA 468
Db 141 ProGlyLysProThrThrSerGluLysIleAsnMetIleSerGly 156

RESULT 3
Q96Q11 PRELIMINARY; PRT; 155 AA.
AC Q96Q11;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DJ54B20.1.1 (Novel SSX family protein (isoform 1)) (Fragment).
GN DJ54B20.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
```



```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9503836; PubMed=7951320;
RA Clark J., Rocques P.J., Crew A.J., Gill S., Shipley J., Chan A.M.,
RA Gusterson B.A., Cooper C.S.;
RT "Identification of a novel gene, SYT and SSX, involved in the t(X;18)
RL (p11.2;q11.2) translocation found in human synovial sarcoma.";
DR Nat. Genet. 7:502-508(1994).
DR EMBL; X79200; CAB36970.1; -.
FT NON_TER 1
SQ SEQUENCE 98 AA; 11136 MW; 1C71F5C8D54513DD CRC64;

Alignment Scores:
Pred. No.: 4.56e-29 Length: 98
Score: 363.00 Matches: 65
Percent Similarity: 91.36% Conservative: 9
Best Local Similarity: 80.25% Mismatches: 7
Query Match: 35.41% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q9Y444 (1-98)
QY 322 TTCGGAAGATCATGCCAAGAACCCAGCAGGAGGAAGAAATGTTTCAAGGAAGTGCCA 381
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 18 TyrAspGlnIleMetProLysLysProAlaGluGluGlyAsnAspSerGluGluValPro 37
QY 382 GAGCANTCTGCCCCACAAATGATGGAAACAGCTGTGCCCCCGGGAATCCAGTACC 441
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 38 GluAlaSerGlyProGlnAsnAspGlyLeuGlyLeuGlySerProGlyLysProThrThr 57
QY 442 TTGAGAGAGATTAAACAGACATCTGGACCCCAAGGGGAAACATGCTGACCCACAGA 501
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 58 SerGluLysIleHisGluArgSerGlyProLysArgGlyGluHisAlaTrpThrHisArg 77
QY 502 CTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATCAGCCACCTCGAGGAAGATGAC 561
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 78 LeuArgGluArgLysGlnLeuValIleTyrGluGluLeuSerAspProGluGluAspAsp 97
QY 562 GAG 564
D : : :
Db 98 Glu 98

RESULT 6
ID Q9WZ9 PRELIMINARY; PRT; 64 AA.
AC Q9WZ9;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DE BA564H18.1 (Synovial sarcoma, X breakpoint 2) (Fragment).
GN SSX2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Whitehead S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL596242; CAD18884.1; -.
FT NON_TER 1
FT NON_TER 64
SQ SEQUENCE 64 AA; 7014 MW; 43ADB72AF2FB9613 CRC64;

Alignment Scores:
Pred. No.: 7.93e-18 Length: 64
Score: 255.00 Matches: 48
Percent Similarity: 84.38% Conservative: 6
Best Local Similarity: 75.00% Mismatches: 10
Query Match: 24.88% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q9Y444 (1-98)
QY 322 TTCGGAAGATCATGCCAAGAACCCAGCAGGAGGAAGAAATGTTTCAAGGAAGTGCCA 381
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 18 TyrAspGlnIleMetProLysLysProAlaGluGluGlyAsnAspSerGluGluValPro 37
QY 382 GAGCANTCTGCCCCACAAATGATGGAAACAGCTGTGCCCCCGGGAATCCAGTACC 441
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 38 GluAlaSerGlyProGlnAsnAspGlyLeuGlyLeuGlySerProGlyLysProThrThr 57
QY 442 TTGAGAGAGATTAAACAGACATCTGGACCCCAAGGGGAAACATGCTGACCCACAGA 501
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 58 SerGluLysIleHisGluArgSerGlyProLysArgGlyGluHisAlaTrpThrHisArg 77
QY 502 CTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATCAGCCACCTCGAGGAAGATGAC 561
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 78 LeuArgGluArgLysGlnLeuValIleTyrGluGluLeuSerAspProGluGluAspAsp 97
QY 562 GAG 564
D : : :
Db 98 Glu 98

RESULT 6
ID Q9WZ9 PRELIMINARY; PRT; 64 AA.
AC Q9WZ9;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DE BA564H18.1 (Synovial sarcoma, X breakpoint 2) (Fragment).
GN SSX2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Whitehead S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL596242; CAD18884.1; -.
FT NON_TER 1
FT NON_TER 64
SQ SEQUENCE 64 AA; 7014 MW; 43ADB72AF2FB9613 CRC64;

Alignment Scores:
Pred. No.: 7.93e-18 Length: 64
Score: 255.00 Matches: 48
Percent Similarity: 84.38% Conservative: 6
Best Local Similarity: 75.00% Mismatches: 10
Query Match: 24.88% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q8WWZ9 (1-64)
QY 283 GAACGTCTCAGATGACTTTTCGGCAGCCTCCAGAGAATCTTCCGAAGATCATGCCCAAG 342
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 GluArgProGlnMetThrPheGlyArgLeuGlnGlySerProLysIleMetProLys 20
QY 343 AAGCAGCAGCAGGAGAAATGTTTGAAGAAAGTGCAGAGGATCTGCCCAACAAAT 402
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 LysProAlaGluGluGlyAsnAspSerGluGluValProGluAlaSerGlyProGluAsn 40
QY 403 GATGGGAACAGCTGTGCCCCCGGGAATCCAAAGTACCTTGGAGAGATTAAACAACA 462
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 41 AspGlyLysGluLeuGlySerProGlyLysProThrThrSerGluLysIleHisGluArg 60
QY 463 TCTGGACCCCAAA 474
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 SerGlyLysArg 64

RESULT 7
Q9NZK4
ID Q9NZK4 PRELIMINARY; PRT; 113 AA.
AC Q9NZK4;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DE SSX2 (Fragment).
GN SSX2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA dos Santos N.R., Torensma R., de Vries T.J., Schreurs M.W.J.,
RA de Bruijn D.R.H., Kater-Baats E., Ruiter D.J., Adema G.J.,
RA van Muijen G.N.P., Geurts van Kessel A.;
RT "Heterogeneous expression of the SSX cancer/testis antigens in human
RT melanoma lesions and cell lines.";
RL Cancer Res. 60:1654-1662(2000).
DR EMBL; AF190791; AAF44724.1; -.
FT NON_TER 1
SQ SEQUENCE 113 AA; 12399 MW; 981EBD852BA31DF8 CRC64;

Alignment Scores:
Pred. No.: 5.29e-17 Length: 113
Score: 247.50 Matches: 52
Percent Similarity: 50.88% Conservative: 6
Best Local Similarity: 45.61% Mismatches: 6
Query Match: 24.15% Indels: 50
DB: 4 Gaps: 1

US-09-975-856-1 (1-576) x Q9NZK4 (1-113)
QY 331 ATCATGCCCAAGAGCCAGCAGCAGGAGAAATGTTTGAAGAGTGCAGAGGATCTCT 390
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 IleMetProLysLysProAlaGluGluGlyAsnAspSerGluGluValProGluAlaSer 20
QY 391 GCGCCCAAAATGATGGGAACACAGCTGTGCCCGGGGAAATCCAAAGTACCTTGGAGAAG 450
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 GlyProGlnAsnAspGlyLysGluLeuGlySerProGlyLysProThrThrSerGluLys 40
QY 451 ATTAACAGACATCTGGA 468
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 41 IleHisGluArgSerGlyAsnArgGluAlaGlnGluLysGluGluArgGlyThrAla 60
QY 468 ----- 468
Db 61 HisArgTrpSerSerGlnAsnThrHisAsnIleGlyArgPheSerLeuSerThrSerMet 80
QY 469 -----CCCAAAAGGGGAAA 483
Db 81 GlyAlaValHisGlyThrProLysThrIleThrHisAsnArgAspPro-LysGlyGlyAs 100

```

QY 484 CATGCTGGACCCACAGACTGCTGTAGAGAAAGCAGCTGG 523
 Db 100 nMetProGlyProThrAspCysValArgGluAsnSerTrp 113
 RESULT 8
 Q80ZT4 PRELIMINARY; PRT; 128 AA.
 ID Q80ZT4 AC Q80ZT4
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to RIKEN cDNA 4930414C09 gene.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN R1
 RP TISSUE=Testis;
 RC Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC048441; AAH48441.1; -;
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001909; KRAB-related.
 DR InterPro; IPR003655; KRAB-related.
 DR Pfam; PF01352; KRAB; 1.
 DR SMART; SM00349; KRAB; 1.
 DR PROSITE; PS0805; KRAB; 1.
 DR PROSITE; PS0806; KRAB-RELATED; 1.
 DR PROSITE; PS0806; KRAB-RELATED; 1.
 SQ SEQUENCE 128 AA; 15290 MW; F88614D1CBFF3B70 CRC64;
 Alignment Scores:
 Pred. No.: 1.5e-13 Length: 128
 Score: 214.50 Matches: 56
 Percent Similarity: 45.00% Conservative: 25
 Best Local Similarity: 31.11% Mismatches: 36
 Query Match: 20.93% Indels: 63
 DB: 11 Gaps: 5
 US-09-975-856-1 (1-576) x Q80ZT4 (1-128)
 QY 25 AGAGAGCCAGGATGATGCTCAATATACAGAGAGTTACGAAAGCGCTTCGATGATTT 84
 Db 9 LysAsnProMetGluValLeuTyrGluProLysAsnIleCysLysAlaPheGlnAsp 28
 QY 85 GCCAAATCTCTCTAAGAAAGAGTGGGAAAGATGAAATCTCGAGAAATCGTCTAT 144
 Db 29 SerThrTyrPheSerAspGluValLeuTyrGlyLysLeuThrGlnTyrGlnLysSerAlaTyr 48
 QY 145 GTGTATATGAAGCTAACTATGAGTCTATGATTAACCTAGGTTCACAGTCAACCTCCCA 204
 Db 49 ValTyrMetLysArgAsnTyrIleArgMetThrAspLeuGlyValThrValAsnGlnPro 68
 QY 205 CTTTTCATCGTAGTAAACGGCTGCAGACTTCACCGGGAATGATTTGGTAACGATCGA 264
 Db 69 ValPheMetArg-----GlyLysGluGln 76
 QY 265 AACACAGGAATCAGGTTGAAGTCTCTCAGATGACTTTTCGGCAGCTCCAGAGATCTTC 324
 Db 77 AspLysGlnSerLeuValGlu----- 83
 QY 325 CCGAGATCATGCCCAAGAGCCACAGAGNAGAAATGTTTGAAGGAAGTCCCGAGAG 384
 Db 84 -----Glylle---GluValHisAsp 89
 QY 385 GCATCTGGCCCAAAAATGATGGGAAACAGCTGTGCCCCCGGGAATCCCAAGTACCTTG 444
 Db 90 Ser----- 90
 QY 445 GAGAGATTAAACAGACATCTGGAACCCAAAAGGGGAAACATGCTGGACCCACAGACTG 504


```
Db 53 GlyGlnGlnTyrGlyGlyTyrArgProThrGlnProGlyProGlnProGlnGln 72
QY 295 ATGACTTTCGGAGCCTCCAGAGAAATCTCCGGAAGATCATGCCCAGAGAGCCAGAGAG 354
Db 73 ArgProTyrGly-----TyrAspGlnIleMetProLysLysProAlaGlu 87
QY 355 GAAGAAATAGTGTGAAGAGAGTCCAGAGGCATCTGCGCCACAAATATGATGGAAACAG 414
Db 88 GluGlyAsnAspSerGluGluValProGluAlaSerGlyProGlnAsnAspGlyLysGlu 107
QY 415 CTGTGCCCCCGGGA 429
Db 108 LeuCysProProGly 112

RESULT 12
Q81ZH0 PRELIMINARY; PRT; 39 AA.
AC Q81ZH0;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE SS18/SSX2 fusion protein (Fragment).
GN SS18/SSX2 FUSION.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wei Y., Sun M., Wang J., Zhu X., Larsson C., Dwight T., Xie Y.,
RA Nilsson G., Larsson O.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV138488; AAN39530.1; -.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4336 MW; 8FC179F66C8C7E0D CRC64;

Alignment Scores:
Pred. No.: 1 51e-05 Length: 39
Score: 137.00 Matches: 25
Percent Similarity: 87.88% Conservative: 4
Best Local Similarity: 75.76% Mismatches: 4
Query Match: 13.37% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q81ZH0 (1-39)
QY 322 TTCCCGAAGATCATGCCAGAGAGCCAGAGAGAAATGTTTGAAGAGAGTGCCA 381
Db 7 TyrAspGlnIleMetProLysLysProAlaGluGlyAsnAspSerGluGluValPro 26
QY 382 GAGGCATCTGGCCACAAATGATGGAAACAGCTGTGC 420
Db 27 GluAlaSerGlyProGlnAsnAspGlyLysGluLeuCys 39

RESULT 13
Q81ZG9 PRELIMINARY; PRT; 39 AA.
AC Q81ZG9;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE SS18/SSX2 fusion protein (Fragment).
GN SS18/SSX2 FUSION.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wei Y., Sun M., Wang J., Zhu X., Larsson C., Dwight T., Xie Y.,
RA Nilsson G., Larsson O.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV138488; AAN39530.1; -.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4336 MW; 8FC179F66C8C7E0D CRC64;

Alignment Scores:
Pred. No.: 1 51e-05 Length: 39
Score: 137.00 Matches: 25
Percent Similarity: 87.88% Conservative: 4
Best Local Similarity: 75.76% Mismatches: 4
Query Match: 13.37% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q81ZH0 (1-39)
QY 322 TTCCCGAAGATCATGCCAGAGAGCCAGAGAGAAATGTTTGAAGAGAGTGCCA 381
Db 7 TyrAspGlnIleMetProLysLysProAlaGluGlyAsnAspSerGluGluValPro 26
QY 382 GAGGCATCTGGCCACAAATGATGGAAACAGCTGTGC 420
Db 27 GluAlaSerGlyProGlnAsnAspGlyLysGluLeuCys 39

RESULT 14
Q81ZG8 PRELIMINARY; PRT; 39 AA.
AC Q81ZG8;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE SS18/SSX2 fusion protein (Fragment).
GN SS18/SSX2 FUSION.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wei Y., Sun M., Wang J., Zhu X., Larsson C., Dwight T., Xie Y.,
RA Nilsson G., Larsson O.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV138490; AAN39532.1; -.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4336 MW; 8FC179F66C8C7E0D CRC64;

Alignment Scores:
Pred. No.: 1 51e-05 Length: 39
Score: 137.00 Matches: 25
Percent Similarity: 87.88% Conservative: 4
Best Local Similarity: 75.76% Mismatches: 4
Query Match: 13.37% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q81ZG8 (1-39)
QY 322 TTCCCGAAGATCATGCCAGAGAGCCAGAGAGAAATGTTTGAAGAGAGTGCCA 381
Db 7 TyrAspGlnIleMetProLysLysProAlaGluGlyAsnAspSerGluGluValPro 26
QY 382 GAGGCATCTGGCCACAAATGATGGAAACAGCTGTGC 420
Db 27 GluAlaSerGlyProGlnAsnAspGlyLysGluLeuCys 39

RESULT 15
Q81ZG7 PRELIMINARY; PRT; 39 AA.
AC Q81ZG7;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE SS18/SSX2 fusion protein (Fragment).
GN SS18/SSX2 FUSION.
```

```
DR EMBL; AV138489; AAN39531.1; -.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4336 MW; 8FC179F66C8C7E0D CRC64;

Alignment Scores:
Pred. No.: 1 51e-05 Length: 39
Score: 137.00 Matches: 25
Percent Similarity: 87.88% Conservative: 4
Best Local Similarity: 75.76% Mismatches: 4
Query Match: 13.37% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q81ZG9 (1-39)
QY 322 TTCCCGAAGATCATGCCAGAGAGCCAGAGAGAAATGTTTGAAGAGAGTGCCA 381
Db 7 TyrAspGlnIleMetProLysLysProAlaGluGlyAsnAspSerGluGluValPro 26
QY 382 GAGGCATCTGGCCACAAATGATGGAAACAGCTGTGC 420
Db 27 GluAlaSerGlyProGlnAsnAspGlyLysGluLeuCys 39

RESULT 14
Q81ZG8 PRELIMINARY; PRT; 39 AA.
AC Q81ZG8;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE SS18/SSX2 fusion protein (Fragment).
GN SS18/SSX2 FUSION.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wei Y., Sun M., Wang J., Zhu X., Larsson C., Dwight T., Xie Y.,
RA Nilsson G., Larsson O.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV138490; AAN39532.1; -.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4336 MW; 8FC179F66C8C7E0D CRC64;

Alignment Scores:
Pred. No.: 1 51e-05 Length: 39
Score: 137.00 Matches: 25
Percent Similarity: 87.88% Conservative: 4
Best Local Similarity: 75.76% Mismatches: 4
Query Match: 13.37% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q81ZG8 (1-39)
QY 322 TTCCCGAAGATCATGCCAGAGAGCCAGAGAGAAATGTTTGAAGAGAGTGCCA 381
Db 7 TyrAspGlnIleMetProLysLysProAlaGluGlyAsnAspSerGluGluValPro 26
QY 382 GAGGCATCTGGCCACAAATGATGGAAACAGCTGTGC 420
Db 27 GluAlaSerGlyProGlnAsnAspGlyLysGluLeuCys 39

RESULT 15
Q81ZG7 PRELIMINARY; PRT; 39 AA.
AC Q81ZG7;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE SS18/SSX2 fusion protein (Fragment).
GN SS18/SSX2 FUSION.
```

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wei Y., Sun M., Wang J., Zhu X., Larsson C., Dwight T., Xie Y.,
 RA Nilsson G., Larsson O.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY138491; AAN39533.1; -;
 FT NON_TER 1
 FT NON_TER 39
 FT NON_TER 39
 SQ SEQUENCE 39 AA; 4336 MW; 8FC179F66C8C7B0D CRC64;

Alignment Scores:
 Pred. No.: 1.51e-05 Length: 39
 Score: 137.00 Matches: 25
 Percent Similarity: 87.88% Conservative: 4
 Best Local Similarity: 75.76% Mismatches: 4
 Query Match: 13.37% Indels: 0
 DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q8IZG7 (1-39)

Qy 322 TTCGGAGATCATGCCCAAGACCCAGACAGAGAGAAATGGTTTCAAGGAAGTGCCA 381
 Db 7 TyrAspGlnIleMetProLysLysProAlaGluGlyAsnAspSerGluGluValPro 26
 Qy 382 GAGGCATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 420
 Db 27 GluAlaSerGlyProGlnAsnAspGlyLysGluLeuCys 39

Search completed: March 31, 2004, 13:59:50
 Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 15:16:19 ; Search time 2745 Seconds
(without alignments)
6266.162 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 576

Sequence: 1 ATGACGAGACACGCCCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_estum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estio:*

8: em_hic:*

9: gb_esti:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_man:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_paq:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	496.6	86.2	979	12	BM806411
2	495	85.9	836	13	BUI194624
3	495	85.9	887	13	BUI161779
4	491	85.2	863	13	BQ222907
					AGENCOURT

5	488.6	84.8	1001	12	BM471266	AGENCOURT
6	481.6	83.6	781	13	BQ434972	AGENCOURT
7	481.6	83.6	873	13	BUI170242	AGENCOURT
8	481.6	83.6	888	13	BQ432374	AGENCOURT
9	479.2	83.2	1135	10	BE535379	AGENCOURT
10	471.6	81.9	602	14	CB158955	601058532
11	470.8	81.7	950	13	BQ231752	AGENCOURT
12	454.6	78.9	793	12	BI560469	603254435
13	447	77.6	922	13	BQ229996	AGENCOURT
14	423.8	73.6	573	10	BE408883	601303758
15	423.4	73.5	867	10	BE891434	601348893
16	417.6	72.5	627	14	CD767542	AGENCOURT
17	408.8	71.0	589	10	BE410950	601303658
18	402	69.8	943	13	BQ229064	AGENCOURT
19	400.8	69.6	948	12	BG333981	602460335
20	394.4	68.5	630	10	BE408978	601303986
21	394.4	68.5	660	14	CB156144	601303986
22	394.4	68.5	759	10	BE384545	601277946
23	394.4	68.5	798	10	BE410794	601301423
24	388.6	67.5	866	12	BG478295	602523210
25	383.4	66.6	636	10	BE390290	601285178
26	383	66.5	678	10	BF211314	601812555
27	382.4	66.4	757	12	BG104299	602310889
28	377.6	65.6	1053	13	BQ224117	AGENCOURT
29	373	64.8	848	10	BE387586	601274054
30	371.4	64.5	961	13	BUI189304	AGENCOURT
31	370.8	64.4	662	10	BE391023	601285631
32	367.6	63.8	1057	12	BG476029	602521175
33	360.2	62.5	655	12	BI831262	603080529
34	360	62.5	888	10	BE729870	601564905
35	360	62.5	1007	12	BG396977	602433891
36	357.6	62.1	1006	13	BUI189362	AGENCOURT
37	340.6	59.1	533	10	BF184274	601843339
38	339.4	58.9	958	13	BQ232177	AGENCOURT
39	330.6	57.4	623	10	BE896093	601438895
40	326	56.6	367	10	BE866727	601679334
41	324	56.2	927	13	BUI178938	AGENCOURT
42	322	55.9	486	9	AA312651	EST183327
43	322	55.9	896	10	BF212335	601813607
44	319.8	55.5	484	14	CB143725	601303197
45	317	55.0	820	10	BE777476	601348520

ALIGNMENTS

RESULT 1	BM806411	979 bp	mRNA	linear	EST 05-MAR-2002
LOCUS	AGENCOURT_6542819	NIH_MGC_72	Homo sapiens	cdna clone	IMAGE:5548810
DEFINITION	5', mRNA sequence.				
ACCESSION	BM806411				
VERSION	BM806411.1	GI:19123234			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 979)				
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: ATCC/DCTD/DTF CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLNL2258 row: f column: 11 High quality sequence stop: 622.				

FEATURES source Location/Qualifiers

1..979

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5548810"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 72"

/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 86.2%; Score 496.6; DB 12; Length 979;

Best Local Similarity 91.5%; Pred. No. 1.9e-131;

Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60

DB 65 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 124

QY 61 TTACGAAAGCCCTTCGATGATATTCGCAAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120

DB 125 ATCCAAAGGCCCTTCGATGATATTCGCAAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 184

QY 121 AAATCCTCGAGAAATCGTGTATGTATATGATGAAGTAACTATGAGTTCATGACTAA 180

DB 185 AAAGCCCTCGAGAAATCTTCTATGTATATGATGAAGAAAGATGAGGCTATGACTTAA 244

QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCTGTAGTAAACGGGCTGCAGACTTCCAC 240

DB 245 CTAGGTTTCAAGGCCACCTTCCACCTTTCATGCTGTAGTAAACGGGCCGAGACTTCCAG 304

QY 241 GGGAATGATTTGGTAAACGATCGAAACCCAGCAATCAGTTGAAAGCTCTTCAGATGACT 300

DB 305 GGGAATGATTTGGTAAATACCTTAACCGTGGGAATCAGTTGAAAGCTCTTCAGATGACT 364

QY 301 TTCCGAGCCCTCAGAGAAATCTTCCGAGATCATGCCGAGAGCCAGCAGAGGAGAA 360

DB 365 TTCCGAGCCCTCAGAGAAATCTTCCGAGATCATGCCGAGAGCCAGCAGAGGAGAA 424

QY 361 AATGGTTTGAAGGAAGTGCAGAGGATCTGGGCCCAAAATATGATGGAAACAGCTGTGC 420

DB 425 AATGATTCGGAGGAGTGCAGAGGATCTGGGCCCAAAATATGATGGAAACAGCTGTGC 484

QY 421 CCCCAGGAAATCGAAGTACCTTGGAGAGATTAACAAAGACATCTGGACCCCAAGGGGG 480

DB 485 CCCCAGGAAATCGAAGTACCTTGGAGAGATTAACAAAGACATCTGGACCCCAAGGGGG 544

QY 481 AAACATGCTGGACCCACAGATCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 540

DB 545 GAACATGCTGGACCCACAGATCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 604

QY 541 AGCGACCTTGAGGAAGATGACGAGTAACTCCCTC 575

DB 605 AGCGACCTTGAGGAAGATGACGAGTAACTCCCTC 639

RESULT 2

LOCUS BU194624

DEFINITION AGENCOURT_7969283 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6165903 5', mRNA sequence.

ACCESSION BU194624

VERSION BU194624.1 GI:22708608

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 836)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/BTP

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13525 row: f column: 16

High quality sequence stop: 624.

FEATURES

source Location/Qualifiers

1..836

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6165903"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 72"

/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 85.9%; Score 495; DB 13; Length 836;

Best Local Similarity 91.3%; Pred. No. 4.9e-131;

Matches 525; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60

DB 74 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 133

QY 61 TTACGAAAGCCCTTCGATGATATTCGCAAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120

DB 134 ATCCAAAGGCCCTTCGATGATATTCGCAAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 193

QY 121 AAATCCTCGAGAAATCGTGTATGTATATGAACTAACTATGAGTTCATGACTTAA 180

DB 194 AAAGCCCTCAGAGAAATCTTCTATGTATATGAAAGAAAGATGATGAGGCTATGACTAA 253

QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCTGTAGTAAACGGGCTGCAGACTTCCAC 240

DB 254 CTAGGTTTCAAGGCCACCTCCACCTTTCATGCTGTATTAACGGGCCGAGACTTCCAG 313

QY 241 GGGAATGATTTGGTAAACGATCGAAACCCAGCAATCAGTTGAAAGCTCTTCAGATGACT 300

DB 314 GGGAATGATTTGGTAAATGACCTTAACCTGGGAATCAGTTGAAAGCTCTTCAGATGACT 373

QY 301 TTCCGAGCCCTCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360

DB 374 TTCCGAGCCCTCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 433

QY 361 AATGGTTTGAAGGAAGTGCAGAGGATCTGGCCCAAAATATGATGGAAACAGCTGTGC 420

DB 434 AATGATTCGGAGGAAGTGCAGAGGATCTGGCCCAAAATATGATGGAAACAGCTGTGC 493

QY 421 CCCCAGGAAATCGAAGTACCTTGGAGAGATTAACAAAGACATCTGGACCCCAAGGGGG 480

DB 494 CCCCAGGAAATCGAAGTACCTTGGAGAGATTAACAAAGACATCTGGACCCCAAGGGGG 553

QY 481 AAACATGCTGGACCCACAGATCGTGGAGAGAGAGAGAGAGAGAGAGAGATC 540

DB 554 GAACATGCTGGACCCACAGATCGTGGAGAGAGAGAGAGAGAGAGAGATC 613

QY 541 AGCGACCTTGAGGAAGATGACGAGTAACTCCCTC 575

DB 614 AGCGACCTTGAGGAAGATGACGAGTAACTCCCTC 648

```

RESULT 3
BU161779
LOCUS
DEFINITION BU161779 887 bp mRNA linear EST 04-SEP-2002
5', mRNA sequence.
ACCESSION BU161779
VERSION BU161779.1 GI:22675689
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 887)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTP/DTP
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13532 row: a column: 22
High quality sequence stop: 739.
FEATURES
source
1..887
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6168477"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
ORIGIN
Query Match 85.9%; Score 495; DB 13; Length 887;
Best Local Similarity 91.3%; Pred. No. 5.1e-131;
Matches 525; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTCAGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db
75 ATGAACGGAGACGACGCTTTCAGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 134
QY 61 TTACGAAAGGCGCTTCGATGATATGTCCTCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
Db
135 ATCCAAAGGCGCTTCGATGATATGTCCTCAATATCTCTTAAGAAAGAGTGGGAAAGATG 194
QY 121 AAATCCTCGGAGAAATCTCTATGATATGAAGCTAAATATGAGGTCTATGACTAA 180
Db
195 AAAGCGCTCAGAGAAATCTCTATGATATGAAGAAAGATGATGAGGTCTATGACTAA 254
QY 181 CTAGGTTTCAAGGTTCACCTCCACCTTCATGCGTAGTAAACGGGCTCCAGCTTCCAC 240
Db
255 CTAGGTTTCAAGGCCACCTCCACCTTCATGATGTAATAAAGCGGCCGAGACTTCCAG 314
QY 241 GGAATGATTTTGGTAACGATGAAACACAGGAATCAGGTTGAACGCTCTCAGATGACT 300
Db
315 GGAATGATTTTGGTAATGACCTAACCGTGGGATCAGGTTGAACGCTCTCAGATGACT 374
QY 301 TTCGGCAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAA 360
Db
375 TTCGGCAGCTCCAGGGAATCTCCCGAAGATCATGCCCAAGAGCCAGCAGAGAGAA 434
QY 361 AATGGTTTGAAGGAAGTCCAGGAGCTGCGCCCAAAATGATGGGAAACAGCTGTGC 420

```

```

Db
435 AATGATTCGGAGGAAGTCCAGAGAGCATCTGGCCCAAAATGATGGAAAGAGCTGTGC 494
QY 421 CCCCCGGGAATCAAGTACTTTGGAGAGATTAAAGACATCTGGACCCAAAGGGGG 480
Db
495 CCCCCGGGAATCAAGTACTCTGAGAGATTCCAGAGATCTGGACCCAAAGGGGG 554
QY 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAAGATC 540
Db
555 GAACATGCTGGACCCACAGACTGCGTGAGAGAAACAGCTGGTGTATGAAGATC 614
QY 541 AGCGACCTCGAGGAAGATGACGATGATGATGATGATGATGATGATGATGATGATG 575
Db
615 AGCGACCTCGAGGAAGATGACGATGATGATGATGATGATGATGATGATGATGATG 649

```

```

RESULT 4
BQ222907
LOCUS
DEFINITION BQ222907 863 bp mRNA linear EST 02-MAY-2002
5', mRNA sequence.
ACCESSION BQ222907
VERSION BQ222907.1 GI:20404307
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 863)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTP/DTP
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13368 row: h column: 17
High quality sequence stop: 625.
FEATURES
source
1..863
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6095680"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
ORIGIN
Query Match 85.2%; Score 491; DB 13; Length 863;
Best Local Similarity 91.2%; Pred. No. 7e-130;
Matches 521; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 5 ACGGAGAGCGCTTTCGAGGAGACCCAGGATGATGCTCAATATCAGAGAAGTAC 64
Db
1 ACGGAGAGCGCTTTCGAGGAGACCCAGGATGATGCTCAATATCAGAGAAGTAC 60
QY 65 GAAAGCGCTTCGATGATATGCTCAATATCTCTTAAGAAAGAGTGGGAAAGATGAAT 124
Db
61 AAGAGCGCTTCGATGATATGCTCAATATCTCTTAAGAAAGAGTGGGAAAGATGAAG 120
QY 125 CCTCGGAGAAATCGTCTATGATATGAAGCTAAATATGAGGTCTATGACTAACTAG 184
Db
121 CCTCAGAGAAATCTCTATGATATGAAGAGAAAGTATGAGGCTATGACTAACTAG 180

```

QY 185 GTTCAAGTCAACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCACGGGA 244
 Db 181 GTTCAAGGCCACCTCCACCTTTTCATGTAATAAAGCGGCCGAGACTTCCAGGGGA 240
 QY 245 ATGATTTGGTAAACCATGAAACCAAGAAATCAGGTTGAAGCTCTCAGATGACTTTCG 304
 Db 241 ATGATTTGGATGAACCTTAACCGTGGGAATCAGGTTGAACCTCTCAGATGACTTTCG 300
 QY 305 GCAGCTCCAGAGATCTTCCGGAAGATCATGCCAAGAACCCAGAGAGAGAGAAATG 364
 Db 301 GCAGCTCCAGGAATCTCCCGAAGATCATGCCAAGAACCCAGAGAGAGAGAAATG 360
 QY 365 GTTCAAGGAAGTGCAGAGGATCTGGCCCAAAATGATGGGAAAACAGCTGTGCCCCC 424
 Db 361 ATTCGGAAGGAAGTGCAGAGGATCTGGCCCAAAATGATGGGAAAACAGCTGTGCCCCC 420
 QY 425 CGGGAATCCCAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGGAAC 484
 Db 421 CGGGAATCCCAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGGAAC 480
 QY 485 ATGCTGGAACCCACAGACTGGTGGAGAAAGCAGCTGGTGGTTTATGAAGAGATCAGCG 544
 Db 481 ATGCTGGAACCCACAGACTGGTGGAGAAAGCAGCTGGTGGTTTATGAAGAGATCAGCG 540
 QY 545 ACCCTGAGAGATGACAGATTAATCCCTC 575
 Db 541 ACCCTGAGAGATGACAGATTAATCCCTC 571

RESULT 5
 BM471266
 LOCUS BM471266 1001 bp mRNA linear EST 05-FEB-2002
 DEFINITION AGENCOURT_6478381 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5563101
 5', mRNA sequence.
 ACCESSION BM471266
 VERSION BM471266.1 GI:18520308
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1001)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM12293 row: i column: 22
 High quality sequence start: 258
 High quality sequence stop: 436.
 Location/Qualifiers
 1..1001
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5563101"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; dr.
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo
 Average insert size 2 kb. Library constructed by Life
 Technologies."

FEATURES

source

RESULT 6
 BM434972
 LOCUS BM434972 781 bp mRNA linear EST 24-MAY-2002
 DEFINITION AGENCOURT_7896851 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6159943
 5', mRNA sequence.
 ACCESSION BM434972
 VERSION BM434972.1 GI:21174048
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 781)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM13509 row: n column: 08
 High quality sequence stop: 634.
 Location/Qualifiers
 1..781

Query Match

84.8%; Score 488.6; DB 12; Length 1001;

ORIGIN

FEATURES
 source

Best Local Similarity 90.6%; Pred. No. 3.7e-129;
 Matches 521; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGCGCTTTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
 Db 78 ATGAACGGAGACGCGCTTTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 137
 QY 61 TTACGAAGGGCTTTCGATGATATTGCCAAATACATTTCTTAAGAAAGATGGGAAAAGATG 120
 Db 138 ATCCAAAAGGCTTTCGATGATATTGCCAAATACATTTCTTAAGAAAGATGGGAAAAGATG 197
 QY 121 AAATCTCCGAGAGAAAATCGTGTATGTATGAAGCTAAACTATGAGCTCATGACTATAA 180
 Db 198 AGAGCTTCAGAGAAAATCTTCTATGTATGAAGAGAAAGTATGAGCTCATGACTATAA 257
 QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCTAGTAAACCGGGTTCAGACTTCCAC 240
 Db 258 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCTAGTAAACCGGGTTCAGACTTCCAC 317
 QY 241 GGGATGATTTTGGTAAAGTCCGAAACACAGGAAATCAGGTTTGAACGTTCCCTCAGATGACT 300
 Db 318 GGGATGATTTTGGTAAAGTCCGAAACACAGGAAATCAGGTTTGAACGTTCCCTCAGATGACT 377
 QY 301 TTCCGAGCGCTCCAGAGAAATCTTCCGAAAGATCATGCCAAGAACCCAGAGAGAGAGAGAA 360
 Db 378 TTCCGAGCGCTCCAGAGAAATCTTCCGAAAGATCATGCCAAGAACCCAGAGAGAGAGAGAA 437
 QY 361 AATGCTTGAAGGAAGTCCAGAGGATCTGGCCCAAAATGATGGGAAAACAGCTGTGC 420
 Db 438 AATGCTTGAAGGAAGTCCAGAGGATCTGGCCCAAAATGATGGGAAAACAGCTGTGC 497
 QY 421 CCCCAGGAAATCCAACTTACCTTGGAGAAATTAACAGACATCTGGACCCAAAGGGG 480
 Db 498 CCCCAGGAAATCCAACTTACCTTGGAGAAATTAACAGACATCTGGACCCAAAGGGG 557
 QY 481 AAACATGCTGGAGAGAGATGACGAGTAAAGCAGCTGGTGGTTTATGAAGAGATC 540
 Db 558 GAACATGCTGGAGAGAGATGACGAGTAAAGCAGCTGGTGGTTTATGAAGAGATC 617
 QY 541 AGCGACCTGAGGAGATGACGAGTAACTCCCTC 575
 Db 618 AGCGACCTGAGGAGATGACGAGTAACTCCCTC 652


```

BQ432374
LOCUS      BQ432374      888 bp      mRNA      linear      EST 24-MAY-2002
DEFINITION AGENCOURT 7859139 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6167925
5', mRNA sequence.
ACCESSION BQ432374
VERSION   BQ432374.1 GI:21171450
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished (1999)
JOURNAL   Contact: Robert Strausberg, Ph.D.
COMMENT   Tissue Procurement: ATCC/DCTD/DTF
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Sequencing by: The I.M.A.G.E. Consortium (ILNL)
           Clone Distribution: Agencourt Bioscience Corporation
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLM13530 row: j column: 22
           High quality sequence stop: 636.
           Location/Qualifiers
             1..888
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:6167925"
               /tissue_type="melanotic melanoma"
               /lab_host="DH10B (phage-resistant)"
               /clone_lib="NIH MGC 72"
               /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
               Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
               Average insert size 2 kb. Library constructed by Life
               Technologies."
FEATURES             source
   Query Match      83.6%; Score 481.6; DB 13; Length 888;
   Best Local Similarity 89.8%; Pred. No. 3.6e-127;
   Matches 517; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGCGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 79 ATGAACGGAGACGACACTTTGCAAGGAGACCCAGGATGATGCTCAAGCATCAGAGAAG 138
QY 61 TTACGAAGGCGCTTCGATGATATTTGCCAATACTCTCTTAAGAAAGAGTGGGAAAGATG 120
DB 139 AGAAGCAAGGCGCTTTGATGATATTTGCCACATCTCTCTTAAGAAAGAGTGGGAAAGATG 198
QY 121 AAATCTCTCGGAGAAATCGCTATGTCTATGAAGCTAACTATGAGTTCATGACTTAA 180
DB 199 AAATCTCTCGGAGAAATCGCTATGTATGATGAAGAACTATAAGGCCATGACTTAA 258
QY 181 CTAGGTTTCAAGTTCACCTCCACCTTTTCATGCGTAGTAAACGGGTCGAGACTTCCAC 240
DB 259 CTAGGTTTCAAGTTCACCTCCACCTTTTCATGCGTAGTAAACGGGTCGAGACTTCCAG 318
QY 241 GGGATGATTTTGGTACGATCGAACCACAGGAATCAGGTTGAAACGCTCTCAGATGACT 300
DB 319 GGGATGATTTTGGTACGATCGAACCACAGGAATCAGGTTGAAACGCTCTCAGATGACT 378
QY 301 TTCCGACGCTCCAGAGAAATCTTCCGAGATCATGCTCCCAAGAGCCAGCAGAGAGAA 360
DB 379 TTCCGACGCTCCAGAGAAATCTTCCGAGATCATGCTCCCAAGAGCCAGCAGAGAGAA 438
QY 361 AATGGTTTGAAGGAAGTGCAGAGGCTATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 420
DB 439 AATGATTCGAAGGAGGTGTCAGAGCATCTGGCCCAACAAACGATGGGAAACAGCTGTGC 498

```

```

QY 421 CCCCCGGGAATCCAACTACCTTGGAGAGATTAAACAAGATCTGGACCCCAAGGGGG 480
DB 499 CCCCCAGGAAAGCAAAATATTTCTGAGAGATTATTAAGAGATCTGGACCCCAAGGGGG 558
QY 481 AAACATGCTCGACCCACAGACTGCTGAGAGAAAGCAGCTGCTGTTTATCAAGAGATC 540
DB 559 AAACATGCTCGACCCACAGACTGCTGAGAGAAAGCAGCTGCTGTTTATCAAGAGATC 618
QY 541 AGCGACCTCGAGGAGAGATGACGAGTAATCCCTCG 576
DB 619 AGCGACCTCGAGGAGAGATGACGAGTAATCCCTCG 654

RESULT 9
LOCUS    BE535379
DEFINITION BE535379 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3445470 5',
mRNA sequence.
ACCESSION BE535379
VERSION   BE535379.1 GI:9764024
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished (1999)
JOURNAL   Contact: Robert Strausberg, Ph.D.
COMMENT   Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: ATCC
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone Distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLM8415 row: o column: 07
           High quality sequence stop: 650.
           Location/Qualifiers
             1..1135
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:3445470"
               /cell_line="MGC36"
               /lab_host="DH10B"
               /clone_lib="NIH MGC 10"
               /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
               Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
               Average insert size 1.5 kb. Library prepared by Life
               Technologies."
FEATURES             source
   Query Match      83.2%; Score 479.2; DB 10; Length 1135;
   Best Local Similarity 89.9%; Pred. No. 2e-126;
   Matches 514; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACCCCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 62 ATGAACGGAGACGACACCTTTGCAAGGAGACCCAGGATGATGCTCAAGCATCAGAGAAG 121
QY 61 TTACGAAGGCGCTTCGATGATATTTGCCAATACTCTCTTAAGAAAGAGTGGGAAAGATG 120
DB 122 AGAAGCAAGGCGCTTTGATGATATTTGCCACATCTCTCTTAAGAAAGAGTGGGAAAGATG 181
QY 121 AAATCTCTCGGAGAAATCGCTATGTCTATGAAGCTAACTATCAGCTCATGACTTAA 180
DB 182 AAATCTCGGAGAAATCGCTATGTATGATGAAGAGAACTATAAGGCCATGACTTAA 241
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCGTAGTAAACGGGTCGAGACTTCCAC 240

```



```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6066127"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN
Query Match      81.7%; Score 470.8; DB 13; Length 950;
Best Local Similarity 89.9%; Pred. No. 4.8e-124;
Matches 516; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

QY 1 ATGAACGAGACGAGCCTTTGCAAGGAGACCCAGGAGTGTCTCAATATCAGAGAAG 60
DB 74 ATGAACGAGACGAGCCTTTGCAAGGAGACCCAGGAGTGTCTCAATATCAGAGAAG 133
QY 61 TTACGAAAGGCGCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGAAGAAGTG 120
DB 134 ATCCAAAGGCGCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGAAGAAGTG 193
QY 121 AAATCCTCGAGAAAATCGTGTATGTATATGAAGCTAAACTATGAGGTCACTATAA 180
DB 194 AAAGCCTCAGAGAAAATCTTCTATGTGTATATGAAGAGAAAGTATGAGGCTATGACTAAA 253
QY 181 CTAGGTTTCAGGTCACCTCCACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
DB 254 CTAGGTTTCAGGTCACCTCCACCTTTCATGCTAGTAAAGCGGCGGGAAGACTTCCAG 313
QY 241 GGGAAATGATTTGGTAAACGATCGAAACCCAGGAAATCAGGTTGAACGTCCTCAGACT 300
DB 314 GGGAAATGATTTGGTAAACGATCGAAACCCAGGAAATCAGGTTGAACGTCCTCAGACT 373
QY 301 TTCGGAGCCTCCAGAGAAATCTTCCGAAATATATGCCAAGAGCCAGCAGAGAGAA 360
DB 374 TTCGGAGCCTCCAGAGAAATCTTCCGAAATATATGCCAAGAGCCAGCAGAGAGAA 433
QY 361 AATGGTTGAAGGAAGTCCAGAGGATCTGGCCCAACAAATATGGAACACGCTGTGC 420
DB 434 AATGATTCGAGGAAGTCCAGAGGATCTGGCCCAACAAATATGGAACACGCTGTGC 493
QY 421 CCCCCGGGAATCCAGTACCTTCGAGAGATTAACAGACATCTGGACCCCAAAAGGGGG 480
DB 494 CCCCCGGGAATCCAGTACCTTCGAGAGATTAACAGACATCTGGACCCCAAAAGGGGG 553
QY 481 AATCAT-GCCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGTGTTTATGAGAGAT 539
DB 554 GAACATGCGCTGGACCCACAGACTGCGTGAGAGAAACAGCTGTGTTTATGAGAGAT 613
QY 540 CAGCAGCCTCGAGGAAGATCAGAGTAACCTCCCC 573
DB 614 CAGCAGCCTCGAGGAAGATCAGAGTAACCTCCCC 647

RESULT 12
BI560469
LOCUS
DEFINITION 603254435F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296880 5',
mRNA sequence.
ACCESSION BI560469
VERSION BI560469.1 GI:15447770
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL

```

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: csapbe@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11751 row: e column: 09
High quality sequence stop: 790.

FEATURES

source

Location/Qualifiers

1..793

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5296880"

/lab_host="DH10B"

/clone_lib="NIH_MGC_97"

/notes="Organ: testis; Vector: pBluescriptR (modified

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI

5'-TTTTTTTTTTTTTNN-3', size-selected for average

insert size 2.2 kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: this is

a NIH_MGC Library."

ORIGIN

```

Query Match      78.9%; Score 454.6; DB 12; Length 793;
Best Local Similarity 90.1%; Pred. No. 2e-119;
Matches 520; Conservative 0; Mismatches 54; Indels 3; Gaps 3;

QY 1 ATGAACGAGACGAGCCTTTGCAA-GGAGACCCAGGAGTGTCTCAATATCAGAGA 59
DB 143 ATGAACGAGACGAGCCTTTGCAAGGAGACCCAGGTTGGTGTCTCAATATCAGAGA 202
QY 60 GTTACGAAGGCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGGAAAGAT 119
DB 203 GATCCAAAAGGCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGGAAAGAT 262
QY 120 GAAATCTCGAGAAAATCGTGTATGTATGAAGCTAAACTATGAGGTCTAGCTAA 179
DB 263 GAAGGCTCGAGAAAATCTCTATGTGTATGAAGAGAAAGTATGAGGCTATGACTAA 322
QY 180 ACTAGGTTTCAAGGTCACCTCCACCTTTTCATGCGTAGTAAACGGGCTCGAGACTTCCA 239
DB 323 ACTAGGTTTCAAGGTCACCTCCACCTTTTCATGCGTAGTAAACGGGCTCGAGACTTCCA 382
QY 240 CGGGAATGATTTGGTAAACGATCGAAACCCAGGAAATCAGGTTGAACGCTCAGATGAC 299
DB 383 -GGGAATGATTTGGTAAACGATCGAAACCCAGGAAATCAGGTTGAACGCTCAGATGAC 441
QY 300 TTTCCGAGCCTCCAGAGATCTTCCGAGAGATCATGCCCAAGCAGCAGCAGAGAGA 359
DB 442 TTTCCGAGCCTCCAGAGATCTTCCGAGAGATCATGCCCAAGCAGCAGCAGAGAGA 501
QY 360 AATGTTTGAAGGAAGTCCAGAGGATCTTGCCCCACAAAATGATGGGAAACAGCTGTG 419
DB 502 AATGTTTGAAGGAAGTCCAGAGGATCTTGCCCCACAAAATGATGGGAAACAGCTGTG 561
QY 420 CCCCCGGGAATCCAGTACCTTCGAGAGATTAACAGACATCTGGACCCCAAAAGGGGG 479
DB 562 CCCCCGGGAATCCAGTACCTTCGAGAGATTAACAGACATCTGGACCCCAAAAGGGGG 621
QY 480 GAAACATGCTCGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTGTTTATGAA-GAGA 538
DB 622 GGAACATGCTCGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTGTTTATGAA-GAGA 681
QY 539 TCAGGACCTCGAGGAAGATGACAGTAACCTCCCTC 575

```



```

Db      682 TCAGCGACTGAGGAAGATGACGAGTAACCTCCCTC 718
|||||
RESULT 13
BQ229996
LOCUS   BQ229996
DEFINITION BQ229996 322 bp mRNA linear EST 02-MAY-2002
5', mRNA sequence.
ACCESSION BQ229996
VERSION   BQ229996.1 GI:20411396
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTP/Drp
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13297 row: g column: 17
High quality sequence stop: 483.

FEATURES
source
Location/Qualifiers
1..922
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6048424"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6, Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 77.6%; Score 447; DB 13; Length 922;
Best Local Similarity 89.1%; Pred. No. 3.3e-117;
Matches 516; Conservative 0; Mismatches 60; Indels 3; Gaps 3;

QY 1 ATGAACGGAGACGACCCCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db      |||||
QY 44 ATGAACGGAGACGACACCTTTGCAAGAGACCCAGGATGATGCTAAAGCATCAGAGAAG 103
Db      |||||
QY 61 TTACGAAAGCGCTTCGATATATGCGAATATCTCTTAAGAAAGAGTGGAAAGATG 120
Db      |||||
QY 104 AGAAGCAAGCGCTTTGATGATATGCGACATATCTCTTAAGAAAGAGTGGAAAGATG 163
Db      |||||
QY 121 AAATCCTCGGAGAAATCTGCTATGCTATGATATGATCAAGCTAACTATGAGTCTACTAA 180
Db      |||||
QY 164 AAATACTCGGAGAAATCAGCTATGCTATGATATGAGAGAACTATTAAGCCCATGACTAA 223
Db      |||||
QY 181 CTAGGTTTCAAGGTACCTCCACCTTTTCGCGTAGTAAACCGGCTGCAGACTTCCAC 240
Db      |||||
QY 224 CTAGGTTTCAAAAGTCACCTCCACCTTTTCATGTGTATTAACAGCCACAGACTTCCAG 283
Db      |||||
QY 241 GGAATGATTTTGGTAACCATCGAACACACAGGAATCAGTTTGAACGCTCCTCAGATGACT 300
Db      |||||
QY 284 GGAATGATTTTGAATGACCATACCGCAGGATTCAGTTTGAACATCCTCAGATGACT 343
Db      |||||
QY 301 TTCGCAGCGCTCCAGAGAAATCTTCCCGAAGATCATGCCCAAGAACCCAGCAGAGAGAA 360
Db      |||||
QY 344 TTCGCAGCGCTCCAGAGAAATCATCCGAGAGATCATGCCCAAGAACCCAGCAGAGAGAA 403
|||||

361 AATGTTTGAAGAGTGCAGAGGCATCTCGCCACAAAATGATGGAAACAGCTGTGC 420
|||||
Db      |||||
404 AATGATTTCGAGGAGTGTTCAGAAAGCATCTGGCCACAAAACGATGGAAACACTGCAC 463
|||||
QY 421 CCCCCGGGAATCAAGTACCTTCGAGAGATTAAACAGACATCTGACCCAAA-GGGG 479
|||||
Db      |||||
464 CCCCCAGAAAGCAAATTAATCTGAGAAGATTAAAGAGATCTGGACCCAAAAGGGGG 523
|||||
QY 480 GAAACATGCTTGGACCCACAGACTCGCTGAGAGAAAGCAGCTGCTGG-TTTATGAAGAGA 538
|||||
Db      |||||
524 GAAACATGCTTGGACCCACAGACTCGCTGAGAGAAAGCAGCTGCTGG-TTTATGAAGAGA 583
|||||
QY 539 TCAG-CGACCTCGAGGAAGATGACGAGTAACCTCCCTCG 576
Db      |||||
584 TCAGCCGACCTCGAGGAAGATGACGAGTAACCTCCCTCG 622
|||||

RESULT 14
BE408883
LOCUS   BE408883
DEFINITION 601303758F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3637982 5',
mRNA sequence.
ACCESSION BE408883
VERSION   BE408883.1 GI:9345333
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM338 row: d column: 15
High quality sequence stop: 571.

FEATURES
Location/Qualifiers
1..573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3637982"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/notes="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected 550bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

Query Match 73.6%; Score 423.8; DB 10; Length 573;

Best Local Similarity 88.9%; Pred. No. 1.2e-110;

Matches 458; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACCCCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60

Db 59 ATGAACGGAGACGACACCTTTGCAAGAGAGACCCAGGATGATGCTAAAGCATCAGAGAAG 118

QY 61 TTACGAAAGCGCTTCGATGATATTGCCAAATCTCTTCTAAGAAAGAGTGGGAAAGATG 120

```

Db 119 AGAAGCAAGCCCTTTCATGATATTGCCACATATCTCTAAGAAAGAGTGGAAGAAGATG 178
Qy 121 AATCTCTCGGAGAAATCTCTATGTATATGAAGTAACTAATGAGTCTATGACTAAA 180
Db 179 AATATCTCGGAGAAATCAGCTATGTATATGAAGAGAACTATAAGGCCATGACTAAA 238
Qy 181 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCGTAGTAAACGGGCTCAGACTCCAC 240
Db 239 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCGTAGTAAACGGGCTCAGACTCCAC 298
Qy 241 GGGATGATTTTGTAAATCAATCGAATCCACAGGAATCAGGTTGAAGTCTCTCAGATGACT 300
Db 299 GGGATGATTTTGTAAATCAATCGAATCCACAGGAATCAGGTTGAAGTCTCTCAGATGACT 358
Qy 301 TTCCGAGGCTCCAGAGATCTTCCGAGATCATGCCAGGATTCAGGTTGAAGTCTCTCAGATGACT 360
Db 359 TTCCGAGGCTCCAGAGATCATGCCAGGATTCAGGTTGAAGTCTCTCAGATGACT 418
Qy 361 AATGTTTGAAGGAGTGCAGAGGATCTGGCCCAACAAATGATGGGAACAGCTGTGC 420
Db 419 AATGATTCGAAGGAGTGTCAAGAGCATCTGGCCCAACAAATGATGGGAACAGCTGTGC 478
Qy 421 CCCCGGGAATCCAGTACCTTGAGAGGATTAACAAGATCTGGACCCCAAAAGGGGG 480
Db 479 CCCCGGGAATCCAGTACCTTGAGAGGATTAACAAGATCTGGACCCCAAAAGGGGG 538
Qy 481 AAACATGCTGGACCCACAGATCGGTGAGAGAA 515
Db 539 AAACATGCTGGACCCACAGATCGGTGAGAGAA 573

```

```

RESULT 15
BE891434
LOCUS 601434893F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3919991 5',
DEFINITION mRNA sequence.
ACCESSION BE891434
VERSION BE891434.1 GI:10350764
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTN/Drp
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9750 row: j column: 24
High quality sequence stop: 641.
FEATURES
Location/Qualifiers
1..867
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3919991"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/notes="Organ; skin; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

```

```

Query Match 73.5%; Score 423.4; DB 10; Length 867;
Best Local Similarity 89.4%; Pred. No. 1.9e-110;
Matches 512; Conservative 0; Mismatches 56; Indels 5; Gaps 5;
Qy 1 ATGAACGGAGACGACGCGCTTTGCAAGGAGAGCCAGGGATGATGCTCAAAATATCAGAGAAG 60
Db 50 ATGAACGGAGACGACGCGCTTTGCAAGGAGAGCCAGGGATGATGCTCAAAATATCAGAGAAG 109
Qy 61 TTACGAAAGGCGCTTCGATGATATGCCAAATACCTTCTTAAGAAAGAGTGGGAAAAGATG 120
Db 110 ATCCAAAGGCGCTTCGATGATATGCCAAATACCTTCTTAAGAAAGAGTGGGAAAAGATG 169
Qy 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCTATGACTAAA 180
Db 170 AAAGCTCAGAGAAATCTTCTATGTATATGAAGAGAAAGTATGAGGCTATGACTAAA 229
Qy 181 CTAGGTTTCAAGGTCACCTCCCACTTTCATGCGTAGTAAACGGGCTGAGACTTCCAC 240
Db 230 CTAGGTTTCAAGGTCACCTCCCACTTTCATGCGTAGTAAACGGGCTGAGACTTCCAC 288
Qy 241 GGGATGATTTTGTAAATCAATCGAATCCACAGGATTCAGGTTGAAGTCTCTCAGATGACT 300
Db 289 GGGATGATTTTGTAAATCAATCGAATCCACAGGATTCAGGTTGAAGTCTCTCAGATGACT 348
Qy 301 TTCCGAGGCTCCAGAGATCTTCCGAGATCATGCCAGGAGCCAGAGAGAGAAAGAA 360
Db 349 TTCCGAGGCTCCAGAGATCTTCCGAGATCATGCCAGGAGCCAGAGAGAGAAAGAA 408
Qy 361 AATGTTTGAAGGAGTGCAGAGGATCTGGCCCAACAAATGATGGGAACAGCTGTGC 420
Db 409 AATGATTCGAAGGAGTGCAGAGGATCTGGCCCAACAAATGATGGGAACAGCTGTGC 468
Qy 421 CCCCGGGAATCCAGTACCTTGAGAGGATTAACAAGATCTGGACCCCAAAAGGGGG 479
Db 469 CCCCGGGAATCCAGTACCTTGAGAGGATTAACAAGATCTGGACCCCAAAAGGGGG 527
Qy 480 GAAACATGCTGGAGCCCAAGATCTGGGTCGAGAGAAAGAGAGAGTCTGGTGTATGAAAGAGAT 539
Db 528 GAAACATGCTGGAGCCCAAGATCTGGGTCGAGAGAAAGAGAGAGTCTGGTGTATGAAAGAGAT 586
Qy 540 CAGGACCTTCGAGAGAGATGACGAGTAACCTCC 572
Db 587 CAGGACCTTCGAGAGAGATGACGAGTAACCTCC 618

```

Search completed: March 31, 2004, 17:25:05
Job time : 2750 secs